

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 14:36:44 ; Search time 49.5417 Seconds
(without alignments)
828.417 Million cell updates/sec

Title: US-10-087-013-11

Perfect score: 1706

Sequence: 1 GNGSGNEISGNCNPKESYPDW.....KKPEYNSKSKDASEYLKDK 308

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

al number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1706	100.0	308	22	AAAB62151
2	894	52.4	311	22	AAAB62150
3	668.5	39.2	2703	16	AAAR70236
4	668.5	39.2	2710	18	AAAW22482
5	668.5	39.2	2710	21	AAAY77904
6	668.5	39.2	3060	18	AAAW22475
7	668.5	39.2	3060	21	AAAY77905
8	600	35.2	407	22	AAAB62148
9	572	33.5	3542	22	AAAB62142
10	521	30.5	793	16	AAAR70234

11	521	30.5	921	18	AAW22480
12	521	30.5	921	21	AAAY77902
13	514	30.1	294	22	AAAB62147
14	444	26.0	351	22	AAAB62149
15	386	22.6	2228	20	AAW93944
16	367.5	21.5	2182	18	AAW22476
17	367.5	21.5	2182	21	AAAY77906
18	342.5	20.1	2197	21	AAAB18352
19	327.5	19.2	1143	23	AAU76759
20	310	18.2	2913	17	AAW00384
21	296	17.4	1726	17	AAW00385
22	268.5	15.7	749	16	AAAR70233
23	268.5	15.7	749	18	AAW22479
24	268.5	15.7	749	21	AAAY77901
25	268.5	15.7	1086	23	AAU76760
26	263.5	15.4	311	21	AAAY77915
27	262.5	15.4	972	23	AAU76761
28	261	15.3	1421	23	AAU76764
29	260	15.2	616	23	AAW05053
30	259	15.2	1028	14	AAAR41044
31	259	15.2	1115	12	AAAR13457
32	259	15.2	1115	18	AAW22478
33	258.5	15.2	1115	21	AAAY77899
34	258.5	15.2	1501	23	AAU76762
35	256.5	15.0	329	23	ABBO7654
36	252	14.8	1061	16	AAAR70231
37	251	14.7	350	23	ABBO7656
38	251	14.7	1435	16	AAAR70232
39	251	14.7	1435	18	AAW22477
40	251	14.7	1435	21	AAAY77900
41	251	14.7	1604	16	AAAR70105
42	251	14.7	1786	14	AAAR41043
43	248.5	14.6	1807	16	AAAR70106
44	248.5	14.6	1807	22	AAAB5697
45	248.5	14.6	2028	22	AAAB5698

ALIGNMENTS

RESULT 1	AAAB62151	standard; peptide: 308 AA.
ID	AAAB62151	standard; peptide: 308 AA.
AC	AAAB62151;	
DT	29-MAY-2001	(first entry)
XX	P. falciparum varCSA polypeptide Itg2-CS2 DBL2.	
DE	FCR3, varCSA protein; chondroitin sulfate A; CSA; var gene; PFEMP1;	
KW	erythrocyte membrane protein 1; parasitized red blood cell; PRBC;	
KW	malaria; protozoacide; Itg2-CS2 DBL2...	
XX	Plasmodium falciparum.	
OS	Plasmodium falciparum.	
XX	WO200116326-A2.	
XX	08-MAR-2001.	
XX	01-SEP-2000; 2000WO-US24195.	
XX	01-SEP-1999; 99US-0152023.	
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
XX	Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;	
XX	Gysin J, Pouvelle B, Fujii N, Smith J;	
XX	WPI: 2001-235109/24.	
XX	Novel FCR3, varCSA protein, useful for modulating parasitized red blood	
XX	cell binding, sequestration and onset of maternal malaria .	

AC AAY77904;
 XX
 DT 13-JUN-2000 (first entry)
 DE P. falciparum Proj3 binding domain polypeptide.
 XX
 KM DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
 KW DABP; Sialic Acid Binding Protein; SAMP; malaria; vaccine; immunisation;
 KM protozoa; Proj3.
 XX
 OS Plasmodium falciparum.
 XX
 PN US5993827-A.
 XX
 PD 30-NOV-1999.
 XX
 PF 07-JUN-1995; 95US-0487826.
 XX
 PR 10-SEP-1993; 93US-0119677.
 (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Sim KL, Chitnis C, Peterson DS, Su X, Wellens TE, Miller LH;
 XX WPI: 2000-194198/17.
 DR N-PSDB; AA298286.
 XX
 PT Isolated protein binding domains from Plasmodium vivax and Plasmodium
 PT falciparum erythrocyte binding proteins useful for vaccinating against
 XX malaria -
 XX
 PS Disclosure; Columns 79-92; 93pp; English.

XX The invention relates to ebl-1 polypeptides that are encoded by the DBL
 CC (Duffy-binding like) gene family. The ebl-1 proteins are substantially
 CC identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid
 CC Binding Protein (SABP), which are soluble proteins that appear in the
 CC culture supernatant after erythrocytes infected with malaria release
 CC merozoites. Immunochemical studies indicate that DABP and SABP are the
 CC respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
 CC and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
 CC used to vaccinate against malaria, especially caused by P. falciparum.
 CC Immunization with the polypeptide provides effective protection against
 CC malaria. The present sequence represents the Proj3 binding domain
 CC polypeptide.

XX Sequence 2710 AA;

Query Match 39.2%; Score 669.5; DB 21; Length 2710;
 Best Local Similarity 40.9%; Pred. No. 8.9e-51;
 Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

QY 1 GNDGSENEISGCPKESYDMDCKKNIDNSHGACMPRRKLC---VRDLTGGEIRKPE 57
 DB 1358 GKNGRTTVGECNPKESYDMDCKKNIDISHGACMPRRKLC---VRLTGGEIRKPE 57
 QY 58 DITLTKFNCAKEHFMHKKYKKNVNAEMLSGKIPFEFRKQMTYTFDFDIFPGID 117
 DB 1418 NLMDATLTKTAALFTFYSWQYKSKNSKSEAKIILDRGLISQFLSMYTFEDYNDICLND 1417
 QY 118 ISSCRYIKDTSQTIKSLG----DOATTEKGDTHIDNNKLLDEWMTTHGPKWMEGLCA 172
 DB 1478 ISKKQ--NDVAKA-KKIKGFESKDSKSPSGISR-----DEWMTTNGPELWKGMLCA 1527
 QY 173 LTNGLSESE-KKNIIDYSTINKINAEKDDCCLEKFAKROFLRMVYVMSDEFRCRRKRL 231
 DB 1528 LTRYVDTDNKRRIKNDYSYDKVNOQNGNPSLEFAAKPOFLRMVMEGEERCAEROKK 1587
 QY 232 EDKVEVCIKAKYIECKKNKSNNSGVKCKEYENITTKGKYQESQEGKF--NTEKROK 289
 DB 1588 ENIKKACKEINSTQOC--NDAKHNQACRAQVEYENKKEFGSGTNNFVLKANVQPO 1645
 QY 290 KPEYNSYSRKD-----ASEYLKDK 308

DB 1646 DPEIKGYEYKDGVPDIOGNEYLLQK 1670

RESULT 6
 ID AAM22475
 XX AAM22475 standard; Protein: 3060 AA.
 XX
 AC AAM22475;
 XX
 DT 12-SEP-1997 (first entry)
 XX
 DE Plasmodium var-7.
 XX
 KM DBL gene family; SABP; sialic acid binding protein; vaccine; therapy;
 KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;
 KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;
 XX Plasmodium.
 XX
 OS Plasmodium vivax.
 OS Plasmodium falciparum.
 XX
 PN W09640766-A2.
 XX
 PD 19-DEC-1996.
 XX
 PF 07-JUN-1996; 96WO-US09508.
 XX
 PR 07-JUN-1995; 95US-0487826.
 XX
 (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;
 XX Wellens TE;
 XX
 DR WPI: 1997-052231/05.
 DR N-PSDB; AAT72882.

XX New malaria vaccines - contains cysteine-rich DBL family protein
 PT binding domains homologous domains of the Duffy and sialic acid
 PT binding proteins
 XX
 PS Claim 8; Page 61-67; 96pp; English.

XX This sequence represents var-7 of Plasmodium. Var-7 belongs to
 CC the Duffy binding like (DBL) family of genes which have homology to the
 CC Duffy antigen binding protein (DABP) and sialic acid binding protein
 CC (SABP) conserved regions (see AAT72889 and AAT72888 respectively). The
 CC var family of genes modulate cytoadherence and antigenic variation of
 CC Plasmodium infected erythrocytes. SABP and the Duffy antigen binding
 CC protein (DABP) are soluble proteins that appear in the culture
 CC supernatant after infected erythrocytes release merozoites. DABP and
 CC SABP mediate the binding of merozoites and schizonts to the erythrocyte
 CC surface. These proteins are necessary for erythrocyte invasion by the
 CC parasite. This sequence can be used in the compositions of the
 CC invention. The compositions are for the treatment and prevention of
 CC malaria, and comprise either a nucleotide sequence or encoded polypeptide
 CC of the var-1, var-2, var-3 or var-7 genes of the DBL gene family, a
 CC family of genes having homology with conserved regions of DABP and SABP.
 CC The compositions are used for the treatment and prevention of malaria.
 CC They are also used in the preparation of vaccines for inducing a
 CC protective immune response in a mammal to Plasmodium merozoites
 CC (especially Plasmodium falciparum or Plasmodium vivax).
 XX

SO Sequence 3060 AA;

Query Match 39.2%; Score 669.5; DB 18; Length 3060;
 Best Local Similarity 40.9%; Pred. No. 1.1e-50;
 Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

QY 1 GNDGSENEISGCPKESYDMDCKKNIDNSHGACMPRRKLC---VNDLTGGEIRKPE 57
 DB 1356 GKNGRTTVGECNPKESYDMDCKKNIDISHGACMPRRKLC---VRLTGGEIRKPE 1415

[illegible]

	RESULT 7
AA777905	
ID	AA777905 standard; Protein; 3060 AA.
XX	
AC	AAV77905;
XX	
DT	13-JUN-2000 (first entry)
XX	
DE	Plasmodium var-7 polypeptide.
XX	
KM	DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
KM	DAB; Sialic Acid Binding Protein; SABB; malaria; vaccine; immunisation;
KW	protozoacidae, var-7.
XX	
OS	Plasmodium sp.
XX	
PN	US5993827-A.
XX	
PD	30-NOV-1999.
XX	
PF	07-JUN-1995; 95US-0487826.
XX	
PR	10-SEP-1993; 93US-0119677.
XX	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.

DR Sim KL, Chitnis C, Peterson DS, Su X, Welles TE, Miller LH;
WPI: 2000-194198/17.
DR N-PSDB: AA296287.
XX
XX Isolated protein binding domains from Plasmodium vivax and Plasmodium
falciparum erythrocyte binding proteins useful for vaccinating against
malaria -
XX
XX Disclosure: Columns 109-124; 93pp; English.
XX
XX The invention relates to ebl-1 polypeptides that are encoded by the DBL
CC (Duffy-binding like) gene family. The ebl-1 proteins are substantially
CC identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid
CC Binding Protein (SABP), which are soluble proteins that appear in the
CC culture supernatant after erythrocytes infected with malaria release
CC merozoites. Immunochemical studies indicate that DABP and SABP are the
CC respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
CC and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
CC used to vaccinate against malaria, especially caused by P. falciparum,
CC immunization with the polypeptide provides effective protection against
CC malaria. The present sequence represents the var-7 polypeptide.
XX
XX Sequence 3060 AA;
S0

	Query Match	39.2%	Score 669.5	DB 21	Length 3060
	Best Local Similarity	40.9%	Pred. No. 1.1e-50		
	Matches 133	Conservative 54	Mismatches 109	Indels 29	Gaps 9
QY	1	GNDGSNEISGCPNPEXSYPDMDCKKNINDNSHSGACMPRRKQIC--VRDLTGGGEIRKPE	57		
		: : : :			
Db	1356	GKNRITVYEGCNPRESYPDMDCKKNINDSHGACMPRRKQICLYIAHESQENIKITDD	1415		
QY	58	DILTFTFNCAKELTFHPAMHKYKKKNVNAENELKSGKIPDEGRKQMYTFEGDFRIFFGTD	117		
		: : : : : : : : : : : : :			
Db	1416	NLMKAFIKTAAAEFTLSMYYKSKNSDEPAKILDLGLIIPQFLBEMMYTFEGDYRIDCLNTD	1475		
QY	118	ISSCRVYIKDPSQTIKSKLG-----DQATTEKGDPIHIDNNKLDQEWMTIHGGKTMEGICA	172		
		: : : : : : : : : : : : :			
Db	1476	ISKKO--NDVAKA-KDKIGKFEKSGKSPSGISR-----QEWMTINGEIWKGMICA	1525		
QY	173	LTNGLISESE--KKNILDOYSYVKNLNNAEKDDCCELEKFAKPOFLMAYVEMSDPFCRERKTL	231		
		: : : : : : : : : : :			
Db	1526	LTKVITVDIDNKRRIKNDYSYKVNQSQNGNSLEEFPAKPPQFLMMIMTEMGEEFCAENQK	1585		
QY	232	EDKVEDVCINAKDYEGKKNMKSNNSCYVAKKEYENYITGKKTYQESQGRK--NTEKRQK	289		
		: : : : : : : : : : : : : : : : : : : :			
Db	1586	ENIKTKDNCENINSTQOC--NDAKHRCQACBAYQVEYENKKKEFGQNNFVLKANVQPO	1643		
QY	290	KPEYNSYSKRD-----ASEYLKQK	308		
		: : : :			
Db	1644	DPEYKGYEYKDGVPQIOGNEYLQK	1668		

RESULT 8
 AAB62148
 ID AAB62148 standard; peptide; 407 AA.
 XX
 AC AAB62148;
 XX
 DT 29-MAY-2001 (first entry)
 XX
 DE P. falciparum varCSA polypeptide A4 DBL4-gamma.
 XX
 KW FCR3.varCSA protein; chondroitin sulfate A; CSA; var gene; PFEKPL;
 KW erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
 KW malaria; protozoacide; A4 DBL4-gamma.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO200116326-A2.
 XX
 PD 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000WO-US24195.
 XX
 PR 01-SEP-1999; 99US-0152023.
 XX
 RA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
 PI Gysin J, Pouvelle B, Fujii N, Smith J;
 XX
 DR WPI: 2001-235109/24.
 XX
 PT Novel FCR3.varCSA protein, useful for modulating parasitized red blood
 PT cell binding, sequestration and onset of maternal malaria -
 XX
 PS Disclosure; Page 72-73; 78pp; English.
 XX
 CC The invention relates to a P. falciparum FCR3.varCSA protein, that is
 CC capable of binding to chondroitin sulfate A (CSA). The var gene and th
 CC corresponding P.falciparum erythrocyte membrane protein 1 (PFEKPL)
 CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
 CC protein and the encoding gene are useful for treating and preventing
 CC maternal malaria or in a patient afflicted at a risk for contracting
 CC maternal malaria or in a patient afflicted with maternal malaria. The
 CC present sequence represents a P. falciparum varCSA polypeptide

CC modulate adhesion of parasitized red blood cell (PRBC) to CSA`the
CC protein and the encoding gene are useful for treating and preventing
CC maternal malaria in a patient identified at a risk for contracting
CC maternal malaria or in a patient afflicted with maternal malaria. The
CC present sequence represents the P. falciparum FCRI3 varCSA protein.
XX Sequence 3542 AA;

SQ Query Match 33.5%; Score 572; DB 22; Length 3542;
Best Local Similarity 39.9%; Pred No.8e-42;
Matches 131; Conservative 54; Mismatches 105; Indels 38; Gaps 14

XX Erythrocyte binding ligand (EBL) family genes were cloned from
 CC P. falciparum chromosome 7 subsegment libraries constructed during
 CC genetic studies of the chloroquine resistance locus. The 4 genes,
 CC EBL-e1 (AA083526), E3ia (AA083527), EBL-e2 (AA083528) and Pro3
 CC (AA083529), encode the proteins given in AAT70233-36, respectively. The
 CC binding domains of such proteins can be expressed e.g. in E. coli,
 CC yeast, mammalian, insect, and in vaccinia virus and adenovirus-infected
 CC cells, and provide protection against P. falciparum.
 XX
 SQ Sequence 793 AA:
 Query Match 30.5%; Score 521; DB 16; Length 793;
 Best Local Similarity 37.9%; Pred. No. 4.2e-38;
 Matches 127; Conservative 46; Mismatches 106; Indels 56; Gaps 13;
 QY 3 DGSNETSGCNPKESYDPMC-KKNIDNSHGACMPRRORLQVRLDTGGELIRKPPDILT 61
 DB 407 NGRTTVGECYRKETYSMTCDSESKIMGOGACIPPRORLCLHYL-----EKIMT 457
 DB 62 K-----FINCAKETHFAHMKYKRD-NVNAEN---ELKSGKIPGFRKMYTFEGDPR 110
 DB 458 NTNELKVAFTKCAAAETFLLMONYKKDKNGNAEDDEKLKGLTFPDRKQMTYTDYR 517
 QY 111 DIFFGTDISCRIRKDTQSTIRSKLGDQATTEKGDTHID-----NKKLQEWMTING 162
 DB 518 DICLGTDISS---KKDTSKGV-----GKVCNIDDVFYKISNSIRYRKSMTWING 564
 QY 163 PKIWEGMLCALTNGLS-----ESEKKNILQDYSYKLNNAEKDDCCLEKFAKPPFLRW 216
 DB 565 PYIWEGLCALSTYDTSLNVNVPETHRKLTGEGNNNEKVIIFGSDSSTLSKFSERPOFLRW 624
 QY 217 YVEMSDPCREKRLKEDKVEDVCIRAKDYEGCKNNKSNNSCVV--CKEYENYITGKKT 273
 DB 625 LTFWGMNPFCKEQK-EYKV--LLAKCKDDVDGDCGCKGKVCACRQCKQYHSMIGIWD 681
 QY 274 QYESQEGKFNTEKR--QKKPEYNSYSKKDASEYLK 306
 DB 682 NYKKQGRYTEVKKIPLYKEDKDKNSDADRYLK 716
 RESULT 11
 AAT22480
 ID AAM22480 standard; Protein; 921 AA.
 AC AAM22480;
 XX 07-OCT-1997 (first entry)
 Plasmodium E3ia.
 XX DBL gene family; SAMP; sialic acid binding protein; vaccine; therapy;
 KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;
 KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;
 KW Plasmodium.
 XX Plasmodium falciparum.
 OS MO9640766-A2.
 PN 19-DEC-1996.
 PD 07-JUN-1996; 96MO-US09508.
 PF 07-JUN-1996; 96MO-US09508.
 PR 07-JUN-1995; 95US-0487826.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;
 PI Wellens TE;
 XX WPI: 1997-052231/05.
 DR N-PSDB; AAT72895.
 DR

XX New malaria vaccines - contains cysteine-rich DBL family protein
 PT binding domains homologous domains of the Duffy and sialic acid
 PT binding proteins
 XX Disclosure; Page 43-45; 96pp; English.
 XX This sequence represents E3ia of Plasmodium. E3ia belongs to
 CC the Duffy binding like (DBL) family of genes which have homology to the
 CC Duffy antigen binding protein (DABP) and sialic acid binding protein
 CC (SABP) conserved regions (see AAT72889 and AAT72888 respectively). The
 CC var family of genes modulate cytoadherence and antigenic variation of
 CC Plasmodium infected erythrocytes. SABP and the Duffy antigen binding
 CC protein (DABP) are soluble proteins that appear in the culture
 CC supernatant after infected erythrocytes release merozoites. DABP and
 CC SABP mediate the binding of merozoites and schizonts to the erythrocyte
 CC surface. These proteins are necessary for erythrocyte invasion by the
 CC parasite. This sequence can be used in the compositions of the
 CC invention, and comprise either a nucleotide sequence or encoded polypeptide
 CC of the var-1, var-2, var-3 or var-7 genes of the DBL gene family.
 CC family of genes having homology with conserved regions of DABP and SABP.
 CC The compositions are used for the treatment and prevention of malaria.
 CC They are also used in the preparation of vaccines for inducing a
 CC protective immune response in a mammal to Plasmodium merozoites
 CC (especially Plasmodium falciparum or Plasmodium vivax).
 XX
 SQ Sequence 921 AA:
 Query Match 30.5%; Score 521; DB 18; Length 921;
 Best Local Similarity 37.9%; Pred. No. 5.2e-38;
 Matches 127; Conservative 46; Mismatches 106; Indels 56; Gaps 13;
 QY 3 DGSNETSGCNPKESYDPMC-KKNIDNSHGACMPRRORLQVRLDTGGELIRKPPDILT 61
 DB 407 NGRTTVGECYRKETYSMTCDSESKIMGOGACIPPRORLCLHYL-----EKIMT 457
 DB 62 K-----FINCAKETHFAHMKYKRD-NVNAEN---ELKSGKIPGFRKMYTFEGDPR 110
 DB 458 NTNELKVAFTKCAAAETFLLMONYKKDKNGNAEDDEKLKGLTFPDRKQMTYTDYR 517
 QY 111 DIFFGTDISCRIRKDTQSTIRSKLGDQATTEKGDTHID-----NKKLQEWMTING 162
 DB 518 DICLGTDISS---KKDTSKGV-----GKVCNIDDVFYKISNSIRYRKSMTWING 564
 QY 163 PKIWEGMLCALTNGLS-----ESEKKNILQDYSYKLNNAEKDDCCLEKFAKPPFLRW 216
 DB 565 PYIWEGLCALSTYDTSLNVNVPETHRKLTGEGNNNEKVIIFGSDSSTLSKFSERPOFLRW 624
 QY 217 YVEMSDPCREKRLKEDKVEDVCIRAKDYEGCKNNKSNNSCVV--CKEYENYITGKKT 273
 DB 625 LTFWGMNPFCKEQK-EYKV--LLAKCKDDVDGDCGCKGKVCACRQCKQYHSMIGIWD 681
 QY 274 QYESQEGKFNTEKR--QKKPEYNSYSKKDASEYLK 306
 DB 682 NYKKQGRYTEVKKIPLYKEDKDKNSDADRYLK 716
 RESULT 12
 AAT77902
 ID AAY77902 standard; Protein; 921 AA.
 AC AAY77902;
 XX 13-JUN-2000 (first entry)
 DE P. falciparum ebl-1 related polypeptide.
 XX DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
 KW DABP; Sialic Acid Binding Protein; SAMP; malaria; vaccine; immunisation;
 KW Protozoacide.
 XX Plasmodium falciparum.
 OS

XX US593827-A.
 XX 30-NOV-1999.
 XX 07-JUN-1995; 9505-0487826.
 XX 10-SEP-1993; 93US-0119677.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Sim KL, Chittis C, Peterson DS, Su X, Wellens TE, Miller LH;
 XX WPI; 2000-194198/17.
 XX DR N-PSDB; AA298284.
 XX Isolated protein binding domains from Plasmodium vivax and Plasmodium
 PT falciparum erythrocyte binding proteins useful for vaccinating against
 PT malaria -

Disclosure: Columns 61-66; 93pp; English.

CC The invention relates to ebl-1 polypeptides that are encoded by the DBL
 CC (Duffy-binding like) gene family. The ebl-1 proteins are substantially
 CC identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid
 CC Binding Protein (SABP), which are soluble proteins that appear in the
 CC culture supernatant after erythrocytes infected with malaria release
 CC merozoites. Immunochemical studies indicate that DABP and SABP are the
 CC respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
 CC and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
 CC used to vaccinate against malaria, especially caused by P. falciparum.
 CC Immunization with the polypeptide provides effective protection against
 CC malaria.

Sequence 921 AA;

Query Match 30.5%; Score 521; DB 21; Length 921;
 Best Local Similarity 37.9%; Pred. No. 5.2e-38;

Matches 127; Conservative 46; Mismatches 106; Indels 56; Gaps 13;

QY 3 DGSNEISGCPNPKESYPMDC-KKNIDNSHGACMPPRKOKCYVDLGGGIRKPEDILT 61
 DB 407 NGRTTGVGCKRYEYSEMTODESKIKKGOMHACIPPRKOKICLHYL-----EKIMT 457
 QY 62 K-----FINCAKETHPFAMHKKYKD-NVNAEN--ELSGKIPGFRKQMYTFGDFR 110
 DB 458 NFNELKYAFIKCAAEFTFLMONTKKDKNGNAEDLDKLGITPEDFKQMFYTFADYR 517
 QY 111 DIFEGTDISSCRVYKTSQTISKSLGDAQTEKGDTHID-----NKKLOEWTHIG 162
 DB 518 DICLGTDISS--KKDTSKGV-----GKVCNIDIVFYFKISINSIRYKRSWETNG 564
 QY 163 KPIWEGMICALTNGLS-----ESEKNIIDYSYNKLNAEKDDCLEFASKPOFLRW 216
 DB 565 PVIWEGMICALSYDTSINNVNPFTHKKLTGEGNNNEFEVITGSSSTLSKFSERPOFLRW 624
 QY 217 YVWSDGFCERKRLKLEKVEDVCIAKADYEGCKNNKSNNSCVKV--CKEYENYITGKKT 273
 DB 625 LTEMGEFCKEOKK-EKKV--LLAKCKDYPDGDGKNGCKVACKQCKOYHSMIGIWD 681
 QY 274 QYESGEGFNTERR--OKRPEVNSYSKKDASEYTK 306
 DB 682 NYKKOKRGTYEVKKIPIYKEDKDKVNSDADRYLK 716

RESULT 13

AA62147 standard; peptide; 294 AA.

AA62147;

29-MAY-2001 (first entry)

XX

DE P. falciparum varCSA polypeptide R29DBL2-gamma.

XX FCR3, varCSA protein; chondroitin sulfate A; CSA; var gene; PFEMP1;
 KW erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
 KW malaria; protozoacide; R29DBL2-gamma.

OS Plasmodium falciparum.

PN WO200116326-A2.

PD 08-MAR-2001.

PE 01-SEP-2000; 2000WO-US24195.

PR 01-SEP-1999; 99US-0152023.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
 PI Gysin J, Pouvelle B, Fujii N, Smith J;
 XX WPI; 2001-235109/24.

PT Novel FCR3, varCSA protein, useful for modulating parasitized red blood
 PT cell binding, sequestration and onset of maternal malaria -
 PS Disclosure: Page 72; 78pp; English.

CC The invention relates to a P. falciparum FCR3, varCSA protein, that is
 CC capable of binding to chondroitin sulfate A (CSA). The var gene and the
 CC corresponding P. falciparum erythrocyte membrane protein 1 (PFEMP1)
 CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
 CC protein and the encoding gene are useful for treating and preventing
 CC maternal malaria in a patient afflicted at a risk for contracting
 CC maternal malaria or in a patient afflicted with maternal malaria. The
 CC present sequence represents a P. falciparum varCSA polypeptide
 CC R29DBL2-gamma.

Sequence 294 AA;

Query Match 30.1%; Score 514; DB 22; Length 294;
 Best Local Similarity 37.4%; Pred. No. 4.7e-38;

Matches 110; Conservative 48; Mismatches 112; Indels 24; Gaps 11;

QY 5 SNEISGCPNPKESYPMDC-KKN-IDNSHGACMPPRKOKCYVDLGGGIRKPEDILT 62
 DB 7 NSGIDMGNKNNKRNKKNEMOCCKNTFVGDNGVCMPPRRKSCICLHNLLEBOTKKNYQLREA 66
 QY 63 FINCAKETHPFAMHKKYKKNVNAENELKSGKIPGFRKQMYTFGDFRIFGTDISSCR 122
 DB 67 FTICAKKETNLMQDKKKNKNEBELKKGKITPEDFMRIMFTTFEDFRDCLNDMG--- 123
 QY 123 YKDSQTISKSLGD--QATTEKGDTHIDNNKLOEWTHIGPKIWEGLCALTNGLS 180
 DB 124 --KVDYK-VKKNININVENNSKRGFKIDP---ENMWNENGPQIWNGLCALIHADPKD 176
 QY 181 EKKNIIDYSYNKLNAEKDD-----CLEFASKPOFLRWYVWSDGFCERKRLKLEKVE 236
 DB 177 STKN-KDNKYIKRYVITLAKRDGSGNMTLSEFKKRKFLEFWEWDYDCKEROKYLTVEA 235
 QY 237 DVCIRAKDYEGCKNNKSNNSCVKCKEYENYITGKTOYESGEGFNTERRKOK 290
 DB 236 STC---KSIDG-GQLKCDRGCKNNKCKDEYKRYKRRKKKEEMNOD-KYYDKRKNK 284

RESULT 14

AA62149 standard; peptide; 351 AA.

AA62149;

29-MAY-2001 (first entry)

XX

DE P. falciparum varCSA polypeptide A4tresDBL3-gamma.
 XX
 XX FCR3 varCSA protein; chondroitin sulfate A; CSA; var gene; PFEMP1;
 KW erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
 KW malaria; protozoacide; A4tresDBL3-gamma.
 XX
 OS Plasmodium falciparum.
 XX
 PN W0200116326-A2.
 XX
 PD 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000MO-US24195.
 XX
 PR 01-SEP-1999; 99US-0152023.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Scherf A, Miller LH, Gamain B, Barnuch DI, Buffet P, Scheidig C;
 Gysin J, Pouvelle B, Fujii N, Smith J;
 DR WPI; 2001-235109/24.
 XX
 PT Novel FCR3 varCSA protein, useful for modulating parasitized red blood
 cell binding, sequestration and onset of maternal malaria -
 XX
 PS Claim 54; Page 73-74; 78pp; English.
 XX
 CC The invention relates to a P. falciparum FCR3 varCSA protein, that is
 CC capable of binding to chondroitin sulfate A (CSA). The var gene and the
 CC corresponding P.falciparum erythrocyte membrane protein 1 (PFEMP1)
 CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
 CC protein and the encoding gene are useful for treating and preventing
 CC maternal malaria in a patient afflicted at a risk for contracting
 CC maternal malaria or in a patient afflicted with maternal malaria. The
 CC present sequence represents a P. falciparum varCSA polypeptide
 CC A4tresDBL3-gamma.
 CC
 XX
 XX Sequence 351 AA:
 Query Match 26.08; Score 444; DB 22; Length 351;
 Best Local Similarity 34.48; Pred. No. 1.2e-31;
 Matches 115; Conservative 46; Mismatches 131; Indels 42; Gaps 12;

AC AAM93944;
 XX
 XX 30-JUN-1999 (first entry)
 XX
 DE P. falciparum PFEMP1 protein.
 XX
 KW Erythrocyte membrane protein; EMP; PFEMP1; malaria; antilococlusional;
 KW glycosamino-glycan-like moiety; antiaggregational; antimalarial;
 KW antigen receptor; infected erythrocyte; rosette formation; blood cell;
 KW capillary occlusion; cerebral malaria; treatment; vaccine; detection;
 KW medicament; parasite; diagnosis; drug screening.
 XX
 OS Plasmodium falciparum.
 XX
 PN W09915557-A1.
 XX
 PD 01-APR-1999.
 XX
 PF 18-SEP-1998; 98MO-SE01675.
 XX
 PR 19-SEP-1997; 97SE-0003386.
 XX
 PA (KARO-) KAROLINSKA INNOVATIONS AB.
 XX
 PI Barragan A, Carlson J, Fernandez V, Qijun C, Wahlgren M;
 DR WPI; 1999-254692/21.
 XX
 XX New isolated malaria polypeptides.
 PT
 PS Claim 4; Page 67-74; 80pp; English.
 XX
 CC This invention describes a novel Plasmodium falciparum erythrocyte
 CC membrane protein (EMP), PFEMP1, which is capable of binding to a
 CC carbohydrate which exhibits at least one negatively charged
 CC glycosamino-glycan (GAG)-like moiety and has antiaggregational,
 CC antilococlusional and antimalarial activity. The carbohydrates of the
 CC invention are capable of acting as receptors for malaria antigens
 CC present on the surfaces of malaria infected erythrocytes, by binding
 CC to these antigens the carbohydrates prevent rosette formation by the
 CC blood cells, this prevents occlusion of capillaries as is seen in
 CC cerebral malaria caused by Plasmodium falciparum. The products of the
 CC invention can be used to treat malaria or to vaccinate against it, or
 CC used to design a model to identify compounds that bind to PFEMP1. The
 CC carbohydrates, polypeptides and antibodies of the invention can be used
 CC as a medicament for dissolving the rosettes formed by erythrocytes
 CC infected by a malaria parasite. The products can also be used for
 CC detection, diagnosis and drug screening.
 CC
 XX
 XX Sequence 2228 AA:
 Query Match 22.68; Score 386; DB 20; Length 2228;
 Best Local Similarity 24.58; Pred. No. 2.6e-25;
 Matches 104; Conservative 59; Mismatches 123; Indels 138; Gaps 14;

```

Db      1156 POKIENPENLWDEANKKPPPOYOYTNVKLIDENSGTSPRTOTQOASSDNPTTLTHFKR 1215
QY      211 POFIRWYVENSDERCRERK-----KLEDKVE--DVCIAKADYEGCKNNKSN----- 255
Db      1216 PTFIRWFEEMWGESFCREKRKRLKOIKYDCKVENGDVGRCSGDGEACDSISTHDYSTVPSF 1275
QY      256 ---SCVAVCKEYENYITGKTQYESQEGKFNTERKQ-KKPEYNSYSKK-----DAS 302
Db      1276 NCPGCGKHCSYRKMIERKKIEFHQSNAYGQOKTDATRRNNGNTFDKEFCCKLETWPDAA 1335
QY      303 EYLK 306
Db      1336 KFLE 1339

```

Search completed: June 20, 2003, 15:02:56
 Job time : 50.5417 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:01:19 ; Search time 16.8255 Seconds

(without alignments)
538.603 Million cell updates/sec

Title: US-10-087-013-11

Perfect score: 1706

Sequence: 1 GNDGSENEISGNCNPKESEYDPM.....KKPEVNSYSKKDASEYLKDK 308

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

All number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	669.5	39.2	2710	2	US-08-568-459A-12	Sequence 12, Appl
2	669.5	39.2	2710	2	US-08-487-826B-12	Sequence 12, Appl
3	669.5	39.2	2710	4	US-09-210-288-12	Sequence 12, Appl
4	669.5	39.2	3060	2	US-08-487-826B-14	Sequence 14, Appl
5	521	30.5	921	2	US-08-568-459A-8	Sequence 8, Appl
6	521	30.5	921	2	US-08-487-826B-8	Sequence 8, Appl
7	521	30.5	921	4	US-09-210-288-8	Sequence 8, Appl
8	367.5	21.5	2182	2	US-08-487-826B-16	Sequence 16, Appl
9	268.5	15.7	749	2	US-08-568-459A-6	Sequence 6, Appl
10	268.5	15.7	749	2	US-08-487-826B-6	Sequence 6, Appl
11	268.5	15.7	749	4	US-09-210-288-6	Sequence 6, Appl
12	263.5	15.4	311	2	US-08-568-459A-21	Sequence 21, Appl
13	263.5	15.4	311	2	US-08-487-826B-33	Sequence 33, Appl
14	263.5	15.4	311	4	US-09-210-288-21	Sequence 21, Appl
15	259	15.2	1115	2	US-08-568-459A-2	Sequence 2, Appl
16	259	15.2	1115	2	US-08-487-826B-2	Sequence 2, Appl
17	259	15.2	1115	4	US-09-210-288-2	Sequence 2, Appl
18	259	15.2	1115	6	5198347-6	Patent No. 5198347
19	251	14.7	1435	2	US-08-568-459A-4	Sequence 4, Appl
20	251	14.7	1435	2	US-08-487-826B-4	Sequence 4, Appl
21	251	14.7	1435	4	US-09-210-288-4	Sequence 4, Appl
22	237.5	13.9	700	2	US-08-568-459A-10	Sequence 10, Appl
23	237.5	13.9	700	2	US-08-487-826B-10	Sequence 10, Appl
24	237.5	13.9	700	4	US-09-210-288-10	Sequence 10, Appl
25	221	13.0	324	2	US-08-568-459A-17	Sequence 17, Appl
26	221	13.0	324	2	US-08-487-826B-29	Sequence 29, Appl
27	221	13.0	324	4	US-09-210-288-17	Sequence 17, Appl

28	178	10.4	197	6	5198347-2	Patent No. 5198347
29	178	10.4	778	6	5198347-4	Patent No. 5198347
30	156	9.1	277	2	US-08-568-459A-15	Sequence 15, Appl
31	156	9.1	277	2	US-08-487-826B-27	Sequence 27, Appl
32	136	9.1	277	4	US-09-210-288-15	Sequence 15, Appl
33	135	7.9	291	2	US-08-568-459A-13	Sequence 13, Appl
34	135	7.9	291	2	US-08-487-826B-25	Sequence 25, Appl
35	135	7.9	291	4	US-09-210-288-13	Sequence 13, Appl
36	127.5	7.5	282	2	US-08-568-459A-16	Sequence 16, Appl
37	127.5	7.5	282	2	US-08-487-826B-28	Sequence 28, Appl
38	127.5	7.5	282	4	US-09-210-288-16	Sequence 16, Appl
39	119.5	7.0	362	2	US-08-568-459A-18	Sequence 18, Appl
40	119.5	7.0	362	2	US-08-487-826B-30	Sequence 30, Appl
41	119.5	7.0	362	4	US-09-210-288-18	Sequence 18, Appl
42	112	6.6	411	2	US-08-568-459A-19	Sequence 19, Appl
43	112	6.6	411	2	US-08-487-826B-31	Sequence 31, Appl
44	112	6.6	411	4	US-09-210-288-19	Sequence 19, Appl
45	109	6.4	2391	2	US-08-446-855A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-568-459A-12
; Sequence 12, Application US/08568459A
; Patent No. 584306
; GENERAL INFORMATION:
; APPLICANT: Slim, Kim L.
; APPLICANT: Chinitz, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhuan
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelson, Ned
; REGISTRATION NUMBER: 29,655
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2710 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
; US-08-568-459A-12
Query Match 39.2%; Score 669.5; DB 2; Length 2710;

Best Local Similarity 40.9%; Pred. No. 6.7e-53;
Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

```

OY 1 GNDGSENEISGCPKESYPDMCKNNIDSHSGACMPRRKRLC--VADLTQGGELRKE 57
Db 1358 GKNQTTTGVGECNPESYPDMCKNNIDSHSGACMPRRKRLCYYIAHESQOTENIKTD 1417
OY 58 DILTFINCAAKETHFAHMYKKKDNVNAENELSGKPIEGEPRKQMYVTFGDFRDIFFCTD 117
Db 1418 NLKDAFIKTAAEFLSMQYKSKNDSEAKILDRGLIPSOFLRSMMTFGYRDLICLNTD 1417
OY 118 ISSCRVYKDTQSOTIKSLG-----DQATTEKGDTHIDNKKLQEWMTIHGPKIWEGLCA 172
Db 1478 ISKKQ--NDVAKA-KDKIGKFFSKDGSKSPSGLSR-----QEWMTIHGPKIWEGLCA 1527
OY 173 LTNLGSESE-KKNILQDYSTNKLNNAEKDDCCLEKPAKPOFLRWYEMSDPCREKRL 231
Db 1528 LTRKYVTDPNKRRIKNDYSYKVNOSGNGPSLEEFAPKPOFLRWMTIEMGEFCAREROK 1587
OY 232 EDKVEDVICAKDYEGCKNNKNSCVKCKEYENITGKKTQYESQEGKF--NTEKRQK 289
Db 1588 ENIKKACNEINSTQOC--NDAKHRCNQCRAVQYEVENKKKEESGOTNNFVLKANVQPO 1645
OY 290 KPEYNSYSKRD-----ASEYLRK 308
Db 1646 DPEYKGYEKDGVQPIQGNELYLQK 1670

```

RESULT 2

US-08-487-826B-12
Sequence 12, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH21.001CPI
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO

ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-487-826B-12

Query Match 39.2%; Score 669.5; DB 2; Length 2710;
Best Local Similarity 40.9%; Pred. No. 6.7e-53;
Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

```

OY 1 GNDGSENEISGCPKESYPDMCKNNIDSHSGACMPRRKRLC--VADLTQGGELRKE 57
Db 1358 GKNQTTTGVGECNPESYPDMCKNNIDSHSGACMPRRKRLCYYIAHESQOTENIKTD 1417
OY 58 DILTFINCAAKETHFAHMYKKKDNVNAENELSGKPIEGEPRKQMYVTFGDFRDIFFCTD 117
Db 1418 NLKDAFIKTAAEFLSMQYKSKNDSEAKILDRGLIPSOFLRSMMTFGYRDLICLNTD 1417
OY 118 ISSCRVYKDTQSOTIKSLG-----DQATTEKGDTHIDNKKLQEWMTIHGPKIWEGLCA 172
Db 1478 ISKKQ--NDVAKA-KDKIGKFFSKDGSKSPSGLSR-----QEWMTIHGPKIWEGLCA 1527
OY 173 LTNLGSESE-KKNILQDYSTNKLNNAEKDDCCLEKPAKPOFLRWYEMSDPCREKRL 231
Db 1528 LTRKYVTDPNKRRIKNDYSYKVNOSGNGPSLEEFAPKPOFLRWMTIEMGEFCAREROK 1587
OY 232 EDKVEDVICAKDYEGCKNNKNSCVKCKEYENITGKKTQYESQEGKF--NTEKRQK 289
Db 1588 ENIKKACNEINSTQOC--NDAKHRCNQCRAVQYEVENKKKEESGOTNNFVLKANVQPO 1645
OY 290 KPEYNSYSKRD-----ASEYLRK 308
Db 1646 DPEYKGYEKDGVQPIQGNELYLQK 1670

```

RESULT 3

US-09-210-288-12
Sequence 12, Application US/09210288
Patent No. 6392026

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH21.1EWV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-0176
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids


```

1  REFERENCE/DOCKET NUMBER:  NH121.001CPI
2  TELECOMMUNICATION INFORMATION:
3  TELEPHONE:  (619) 235-8550
4  TELEFAX:  (619) 235-0176
5  INFORMATION FOR SEQ ID NO:  8:
6  SEQUENCE CHARACTERISTICS:
7  LENGTH:  921 amino acids
8  TYPE:  amino acid
9  STRANDEDNESS:  single
10 TOPOLOGY:  linear
11 MOLECULE TYPE:  protein
12 HYPOTHETICAL:  NO
13 ORIGINAL SOURCE:
14 ORGANISM:  Plasmodium falciparum
15
16 US-08-568-459A-8

```

Query Match	30.5%;	Score 521;	DB 2;	Length 921;
Best Local Similarity	37.9%;	Pred. No. 8.7e-40;		
Matches 127;	Conservative 46;	Mismatches 106;	Indels 56;	Gaps 13

RESULT 6
US-08-487-826B-8
; Sequence 8, Application US/08487826E
Patent No. 5993827

1 FLING DATE: 10-SEP-1993
 2 CLASSIFICATION: 435
 3 ATTORNEY/AGENT INFORMATION:
 4 NAME: Israel, Ned
 5 REGISTRATION NUMBER: 29,655
 6 REFERENCE/DOCKET NUMBER: NH121.001CP1
 7 TELECOMMUNICATION INFORMATION:
 8 TELEPHONE: (619) 235-8550
 9 TELEFAX: (619) 235-0176
 10 INFORMATION FOR SEQ ID NO: 8:
 11 SEQUENCE CHARACTERISTICS:
 12 LENGTH: 921 amino acids
 13 TYPE: amino acid
 14 STRANDEDNESS: single
 15 TOPOLOGY: linear
 16 MOLECULE TYPE: protein
 17 HYPOTHETICAL: NO
 18 ORIGINAL SOURCE:
 19 ORGANISM: Plasmodium falciparum
 20 OS-08-487-8205-8

Query Match	30.5%;	Score 521;	DB 2;	Length 921;
Best Local Similarity	37.9%;	Pred. No. 8.7e-40;		
Matches 127;	Conservative 46;	Mismatches 106;	Indels 56;	Gaps 13;

Qy	274 QYESQEGKFNTEKR--QKKPEYNSISKDASEYLK 306 : :: : : :
Dp	682 NYKKOKGRYTEVKRIPLYKEDKDVRKSDDARDYLK 716 :: : :: :

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBR: US/08/487,826B

```

```

ADDRESS: Knodde Martens Olson & Bear
STREET: 620 Newport Center Drive
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FWV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 921 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-09-210-288-8

Query Match 30.5%; Score 521; DB 4; Length 921;
Best Local Similarity 37.9%; Pred. No. 8.7e-40;
Matches 127; Conservative 46; Mismatches 106; Indels 56; Gaps 13;

QY 3 DGSNEISGNCNPKESYPDMDC--KKNIDNSHSGCAMPRRROKLCVRLDITOGGEIRRPEDILT 61
DB 407 NGRTTGVGCYKREYSEWTKDESKIKMGQHGACIPRRQKLCIHL-----EKIMT 457
QY 62 K-----FTICAKETHFAHMKYKRD-NVAEN---ELKSGKIPEGRKQMTTFGDER 110
DB 458 NTNELKAFIKCAAEFTLLQWQNKGNNAEDLDEKLGIIPEDEKRMFTFYADYR 517
QY 111 DIFGTDISSGRTIKDTSQTSKSLGQDATTEKGDTHID-----NKLQEWMTIHG 162
DB 518 DICGTDISS--KKDTSKGV-----GKVCNIDVDFYKINSIRTRKRSWMEING 564
QY 163 PKINEGMCALITNGLS-----ESEKNILQDYSYNNKLNNAEKDCCLEKFAKRPQLRW 216
DB 565 PVIWEGMCALYSTDTSLNNVNPETHKRLTEGNNNEKVIFFGSDSTLTSKFSERQFLWR 624
QY 217 YVENSDEGRKRLKLEDEKVEDVCIKADYEGCKNNKSNKSVYK---CKEYENITGKKT 273
DB 625 LTEWGENEFCBOKR-EYKV--LLAKCKDCDVGDKCKGCKVACAKDCKQYHWSWIGWID 681
QY 274 QYESQEGKFNTEKR--OKKPEYNSYSKKDASEYLK 306
DB 682 NYKKQKGRYTEVKKIPLAKEDKDVKNSDARDYIK 716

RESULT 8
US-08-487-826B-16
Sequence 16, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhuan
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach

STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2182 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-487-826B-16

Query Match 21.5%; Score 367.5; DB 2; Length 2182;
Best Local Similarity 25.4%; Pred. No. 4.7e-25;
Matches 107; Conservative 60; Mismatches 120; Indels 127; Gaps 16;

QY 10 GGNPKESYPDMDC--KKNIDNSHSG---ACMPRRROKLCVRLDITOGGE----- 52
DB 874 GPGKEKPPNKKCVTPSGVSTATSGDKGALCVPPRRRLYVGGLSQWASRGDETFEYSS 933
QY 53 -----IRKPEDILTRFINCAKETHFAHMKYKRD-----N 83
DB 934 EATSPASOSESEKRLTAIESAAIETFFLWHKYEKKPPATODGAGIGVLEPPSPPE 993
QY 84 NAENEL-KSGKIPGFRKQMTYTFGDFRDIFF-GTDISSCHYIKDTSQTSKSLGQDAT 141
DB 994 DPQTOLOQTVIIPDPLRQMTYTLADYKDIYSSGNDTSQDTGQTSSNDNLKNLYLE 1093
QY 142 EKGDTHTDNDK-----KLOEWMTIHGPKIWEGLCA 172
DB 1054 ASGSTOEKREKMKQIQAKIKILNGATSGVPPVTKNSVKTPOQTWMEKILANDINNAVCA 1113
QY 173 LT-----NGIS-----ESEKNILQDYSYNNKLNNAEKDCC----- 202
DB 1114 LTYKENDARGTSARIEONKDLKALWDEANKRTEIEKYQYNNVRLDESGAKSDTIOPP 1173
QY 203 CLKFKFASKPQFLRWYVENSDEFCREKRLKLEDEKVEDVCIK---AKDYEG---CKNNKS-- 253
DB 1234 YNVLDLSSSCACRKLRYKTIKKTEYERKQKAYEQKSNYENEGDKCOTQSNNNAN 1293
QY 1174 TLKNFVELPTFRHLHENGNSFCFERAKRLAQIHHECDDEGEKQYSDGEYCEIFRSQ 1293
QY 254 -----NNSCVKCKEYENITGKKTQYESQEGKFNTEK---RQKPEYNSYSKKDAS 302
DB 303 EY 304
DB 1294 EF 1295

RESULT 9
US-08-568-459A-6
Sequence 6, Application US/08568459A
Patent No. 5849306
GENERAL INFORMATION:

Query Match	15.7%	Score 268.5,	DB 2,	Length 749,
Best Local Similarity	26.8%	Pred. No. 1.6e-16;		
Matches	85;	Conservative	45;	Mismatches 116;
			Indels	71,
			Gaps	15

Db 352 NNLYNKFEDSKAYLRSE 368

Query Match	15.7%;	Score 268.5;	DB 2;	Length 749;
Best Local Similarity	26.8%;	Pred. NO. 1.6e-16;		
Matches 85;	Conservative 45;	Mismatches 116;	Indels 71;	Gaps 15

```

Db : 301 KOLEKIC --- ENKNCSEK --- CKNACSSYEKWKERKNEYNLQSKKFDSDKLNK - K 351

```

OY 293 YNSTSK-KDASEYLKDK 308
Db 352 NNLNFKEDSKAYLRSE 368

RESULT 11

US-09-210-288-6
; Sequence 6, Application US/09210288
; Patent No. 6392026
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chlitis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,288
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 749 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
US-09-210-288-6

Query Match 15.7%, Score 268.5; DB 4; Length 749;
Best Local Similarity 26.8%; Pred. No. 1.6e-16;
Matches 85; Conservative 45; Mismatches 116; Indels 71; Gaps 15;

OY 6 NEISGCMRKES--YPMDC-KKIDNSHSGCMPPRQKLC-----VRDLTGGELIRKP 56
Db 109 SSVGCTKTKISKVKKKNNKYNKVTREGVCGPBRQOLGLGYIFILRDNEG- 163
OY 57 EDILTKFNCIAKETHFAWH-KYKKDYNVAENELKSGIPGFRKROMYTFGDFRDPFG 115
Db 164 ---LKDHIINKAA--NYEAMHAKKEYENANGDKICNA-----ILGSTADIGDIYRG 208
OY 116 TDI--SSCRYIKDTSOTIKSLGDAQATTEKGDTHIDNKKLOEMWTIHGPKIWEGMICA 172
Db 209 LDVMDINTNKLSEKQKIFWGGGNSRKQ-----NDNNERNKWMKEQRULIMSSMV-- 260
OY 173 LTNGISESEKNNILODYSYNNLNAAEKDCLEKFAKSPQFLRMVTVENSDFECRRKLE 232
Db 261 -----KHIPGKTCKRRHNNFEK-----LPQFLRMLEKWEDEFCEMGTEV 300

OY 233 DKVEDVCIAKADYEGCKNNKNSNCVCKEYENYITGKRTQYESQEGKENTEKROKPE 292
Db 301 KOLEKIC-----ENKNCSEKK-----CKNACSEYEMKIERKNNENYLOSKKIDSDKLLK-K 351
OY 293 YNSTSK-KDASEYLKDK 308
Db 352 NNLNFKEDSKAYLRSE 368

RESULT 12

US-08-568-459A-21
; Sequence 21, Application US/08568459A
; Patent No. 5849306
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chlitis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelson, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-568-459A-21

Query Match 15.4%, Score 263.5; DB 2; Length 311;
Best Local Similarity 26.3%; Pred. No. 1.4e-16;
Matches 67; Conservative 19; Mismatches 158; Indels 11; Gaps 4;

OY 22 CKKIDNSHSGACMPPRQKLC---VRDLTGGELIRKREDILTFINCAAKETHFAWHKY 78
Db 12 CXXXXXXXACACMPBRQKLCLYYIXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 71
OY 79 KKDYNVAENELKSGIPGFRKROMYTFGDFRDIFFGDISSCRYIKDTSOTIKSLGDO 138
Db 72 XXXXXXXXXXXXXXXXXXXXQFLRSMATYTGDIRDCLNTDIS-----KQNDYXXXXXXX 126
OY 139 ATTEKGDTHIDNKKLOEMWTIHGPKIWEGMICAL-TNGISESEKNNILODYSYNNLNA 197

Db 127 XXXXXXXXXXXXSPGSLRQEWMTNGPEIRWGMICALXXXXXXXXXXXXXXXXXXXXX 186
QY 198 EKDDCCLEFASKPOFLRWYVWMSDFECREKRLKEDKVEDVCIRAKDEGCKNNKNSNC 257
Db 187 XXXXXXXXXXXXSKPSGSLRQEWMTNGPEIRWGMICALXXXXXXXXXXXXXXXXXXXXX 244
QY 258 VKVCKEYENYITGKK 272
Db 245 NOACRAYOYEYENKK 259

RESULT 13
US-08-487-826B-33
; Sequence 33, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbé Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487, 826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29, 655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-487-826B-33

Query Match 15.4%; Score 263.5; DB 2; Length 311;
Best Local Similarity 26.3%; Pred. No. 1.4e-16;
Matches 67; Conservative 19; Mismatches 158; Indels 11; Gaps 4;

Db 12 CXXXXXXXACMPRRQKLC---VRDLTGGEIRKPEDILTFKFINCAKETHEFAWHKY 71
QY 22 CKKNIDNSHSGACMPRRQKLC---VRDLTGGEIRKPEDILTFKFINCAKETHEFAWHKY 78
Db 12 CXXXXXXXACMPRRQKLC---VRDLTGGEIRKPEDILTFKFINCAKETHEFAWHKY 71
QY 79 KKDNNVNAENELKSGRIPEGFRKQMYTTFGDFRIDFPGTDISSCRITIKDTISQTIKSKLGDQ 138
Db 72 XXXXXXXXXXXXSKPSGSLRQEWMTNGPEIRWGMICALXXXXXXXXXXXXXXXXXXXXX 126

QY 139 ATTEGDPHIDNNKLOEWMTLHGPKIWEGMLCAL-TGSLSESEKKNILQDYSYNKLNA 197
Db 127 XXXXXXXXXXXXSKPSGSLRQEWMTNGPEIRWGMICALXXXXXXXXXXXXXXXXXXXXX 186
QY 198 EKDDCCLEFASKPOFLRWYVWMSDFECREKRLKEDKVEDVCIRAKDEGCKNNKNSNC 257
Db 187 XXXXXXXXXXXXSKPSGSLRQEWMTNGPEIRWGMICALXXXXXXXXXXXXXXXXXXXXX 244
QY 258 VKVCKEYENYITGKK 272
Db 245 NOACRAYOYEYENKK 259

RESULT 14
US-09-210-288-21
; Sequence 21, Application US/09210288
; Patent No. 6392026
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbé Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210, 288
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36, 516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-09-210-288-21

Query Match 15.4%; Score 263.5; DB 4; Length 311;
Best Local Similarity 26.3%; Pred. No. 1.4e-16;
Matches 67; Conservative 19; Mismatches 158; Indels 11; Gaps 4;

Db 12 CXXXXXXXACMPRRQKLC---VRDLTGGEIRKPEDILTFKFINCAKETHEFAWHKY 71
QY 22 CKKNIDNSHSGACMPRRQKLC---VRDLTGGEIRKPEDILTFKFINCAKETHEFAWHKY 78
Db 12 CXXXXXXXACMPRRQKLC---VRDLTGGEIRKPEDILTFKFINCAKETHEFAWHKY 71
QY 79 KKDNNVNAENELKSGRIPEGFRKQMYTTFGDFRIDFPGTDISSCRITIKDTISQTIKSKLGDQ 138

[illegible]

RESULT 15

US-08-568-459A-2
; Sequence 2, Application US/08568459A
tent No. 5849306
BUREAU OF INFORMATION

GENERAL INFORMATION:

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: PatentIn Release #1.0, Version #1.25
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/08/568,459A
8  FILING DATE: 07-DEC-1995
9  CLASSIFICATION: 435
10 ATTORNEY/AGENT INFORMATION:
11 NAME: Israel, I. N.
12 REGISTRATION NUMBER: 29,655
13 REFERENCE/DOCKET NUMBER: NIH21.001CPI
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: (619) 235-8550
16 TELEFAX: (619) 235-0176
17 INFORMATION FOR SEQ ID NO: 2:
18 SEQUENCE CHARACTERISTICS:
19 LENGTH: 1115 amino acids
20 TYPE: amino acid
21 STRANDEDNESS: single
22 TOPOLOGY: linear
23 MOLECULE TYPE: protein
24 HYPOTHEICAL: NO
25 ORIGINAL SOURCE:
26 ORGANISM: Plasmodium vivax
27 US-08-568-459A-2

```

Query Match	15.2%	Score 259;	DB 2;	Length 1115;
Best Local Similarity	24.9%	Pred. No. 2.1e-15;		
Matches 89; Conservative		38; Mismatches 127;	Indels 104;	Gaps 14

```
QY      1 GNDGSEIISGNP-----KESYPDMDCCKNIINS 29
          || | : ||| |
Db      223 GNSRRKSSNGCNPPYDIDHKKTISSALINHAFLQNTVMKNCNKKRKRERDWDC-----NT 277
```

QY 30 HSGACMPRRKQICAVYDL--QGGEIRKPEIL-----TKFNCAKETHFAM---- 75
 Db 278 KQVCIPIRRRYQIQCKMELTNLVNNTDHFNDHDIITFRKILYLRKLLIYDAAEGDILLKLN 337
 QY 76 HRYKKDNVAENELKSGKIPGEFRKQMYTPGDFNDIEFGDIDISCRYIKDTSQIKSKL 135
 Db 338 YRYNKD-----FCKDIRSLSDFDIIIMGTDMEGISGVENNLISIF 381
 QY 136 GQATTEKGQDTHIDDKKLOEWMTHIHPRIWGMICALTNLSSEKKNIIQDVSYKLN 195
 Db 382 G---TDEKA-----QQRKQMMNESKADIMTAMYSKYL-----KGNFIWICKLVAV 428
 QY 196 NAEKDDCCLEFASKPOPLRWYVEMSEDFECREKKLIEDKVEDVC--TKADYEGCKNNK 252
 Db 429 NIE-----FOIRWIRENGRDVSELPREYOKLEKCDGKINTYDKRVCK-- 473
 QY 253 SNNSCVKCYEYENITFGKTOYESQEGF-----TTEROKKPEVNS--KKNASY 304
 Db 474 -VPCQNAKCSIDQMTITRRKNQNDVLSNKFISYKAAEVQTAGIYTPIDILKQLEDF 530

Search completed: June 20, 2003, 15:07:04
Job time : 18.8255 secs

DEF 530

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:05:30 ; Search time 32.2489 Seconds
(without alignments)
1033.453 Million cell updates/sec

Title: US-10-087-013-11

Perfect score: 1706
Sequence: 1 GNGSNEISGCPKESYPDW.....KKPEYNSYSKDASEYLKDK 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

1 number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PC7_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/PC7US_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	669.5	39.2	2710	9 US-10-153-273-12	Sequence 12, Appl
2	521	30.5	921	9 US-10-153-273-8	Sequence 8, Appl
3	327.5	19.2	1143	10 US-09-924-154-14	Sequence 14, Appl
4	268.5	15.7	749	9 US-10-153-273-6	Sequence 6, Appl
5	268.5	15.7	1086	10 US-09-924-154-15	Sequence 15, Appl
6	263.5	15.4	311	9 US-10-153-273-21	Sequence 21, Appl
7	262.5	15.4	972	10 US-09-924-154-16	Sequence 16, Appl
8	261	15.3	1421	10 US-09-924-154-13	Sequence 13, Appl
9	259	15.2	1115	9 US-10-153-273-2	Sequence 2, Appl
10	258.5	15.2	1501	10 US-09-924-154-17	Sequence 17, Appl
11	251	14.7	1435	9 US-10-153-273-4	Sequence 4, Appl
12	237.5	13.9	700	9 US-10-153-273-10	Sequence 10, Appl
13	221	13.0	324	9 US-10-153-273-17	Sequence 17, Appl
14	156	9.1	277	9 US-10-153-273-15	Sequence 15, Appl
15	135	7.9	291	9 US-10-153-273-13	Sequence 13, Appl
16	127.5	7.5	282	9 US-10-153-273-16	Sequence 16, Appl
17	119.5	7.0	362	9 US-10-153-273-18	Sequence 18, Appl
18	112	6.6	411	9 US-10-153-273-19	Sequence 19, Appl
19	109	6.4	441	9 US-10-153-668-102	Sequence 102, App

20	109	6.4	634	9 US-10-153-668-98	Sequence 98, Appl
21	109	6.4	634	9 US-10-153-668-100	Sequence 100, App
22	109	6.4	1118	9 US-10-153-668-104	Sequence 104, App
23	102	6.0	1338	10 US-09-402-100-4	Sequence 4, Appl
24	101.5	5.9	271	9 US-10-153-273-14	Sequence 14, Appl
25	101.5	5.9	497	9 US-09-820-843A-32	Sequence 32, Appl
26	101.5	5.9	665	9 US-09-820-843A-107	Sequence 107, App
27	99.5	5.8	411	9 US-10-153-273-20	Sequence 20, Appl
28	97	5.7	380	10 US-09-134-333-13	Sequence 13, Appl
29	94.5	5.5	525	9 US-09-372-348-16	Sequence 16, Appl
30	93.5	5.5	368	10 US-09-925-300-1356	Sequence 1356, Ap
31	93	5.5	743	9 US-10-087-464-53	Sequence 53, Appl
32	92.5	5.5	2789	10 US-09-801-574-57	Sequence 57, Appl
33	92.5	5.4	624	10 US-09-815-242-5391	Sequence 5391, Ap
34	92.5	5.4	627	10 US-09-815-242-5391	Sequence 12305, A
35	92.5	5.4	698	9 US-10-281-478-2	Sequence 2, Appl
36	92.5	5.4	698	10 US-09-726-968-4	Sequence 4, Appl
37	92.5	5.4	824	9 US-10-281-478-1	Sequence 1, Appl
38	92.5	5.4	824	10 US-09-726-968-2	Sequence 2, Appl
39	92	5.4	428	9 US-09-742-096-3	Sequence 46, Appl
40	92	5.4	1786	9 US-09-824-574-7	Sequence 3, Appl
41	92	5.4	2476	9 US-10-223-085-336	Sequence 7, Appl
42	91.5	5.4	776	9 US-10-223-084-336	Sequence 336, App
43	91.5	5.4	776	9 US-10-223-084-336	Sequence 336, App
44	91.5	5.4	776	9 US-10-223-088-336	Sequence 336, App
45	91.5	5.4	776	9 US-10-223-090-336	Sequence 336, App

ALIGNMENTS

RESULT 1
US-10-153-273-12
Sequence 12, Application US/10153273
Patent No. US20020169305A1
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
Chilnis, Chetan
Miller, Louis H.
Peterson, David S.
Su, Xin-zhaun
Welliams, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/153,273
FILING DATE: 21-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:

LENGTH: 2710 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium falciparum
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:
 US-10-153-273-12

Query Match 39.2%; Score 669.5; DB 9; Length 2710;
 Best Local Similarity 40.9%; Pred. No. 8.1e-48;
 Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

QY 1 GNDGSEISGNCNPKESYPDMDC--KKNIDNSHSGACMPRRKOKLC---VRDLTGGEIRKRE 57
 DB 1358 GKNGRTTVEGCNCKESYPMDC--KKNNIDSHSGACMPRRKOKLCYYIAHESQTEENIKTD 1417
 58 DILTKFNCACAKTHFAHMKYKKNVNAENELKSGKIPGFRKQMYTTFGDFRDIFFGTD 117
 1418 NIKDAEIKTAAAEPLSMOYKSKNDSEAKILDRGLIPSOFLSMYTFGDYRDICLNTD 1477
 QY 118 ISSCRYIKDTISOTIKSLG----DOATEKGDTHTDDKKKIDEMWTITGPKTWESMLCA 172
 DB 1478 ISKQ--NDYAKA-KKIGKIFESKODSKSPGSLR-----QEWKKTGPELWKMGLCA 1527
 QY 173 LINGLSESE-KNNIIDYSYNKLNNAEKDDCCLEKFPASRPOFLRWYENSDFECPRRKT 231
 DB 1528 LKRYVYDITNKRIRKNDYSYKVNQNGNPISLEFAAPQFLRWMIEMGEFCAROK 1587
 QY 232 EKVVEDYICAKDYESCKNNKSNNSCVKCKEYENITGKKTQYESQESKFT--NTEKKRK 289
 DB 1588 ENIADKAEINSTQOC--NDAKHRCNQACRAVOEYVENKKKEFSQTNFVLKANVQPO 1645
 QY 290 KPEYNSYSKRD-----ASEYLKDK 308
 DB 1646 DEYKGEYKDGVOPIQGNELYLQK 1670

RESULT 2

US-10-153-273-8
 Sequence 8, Application US/10153273
 Patent No. US20020169305A1
 GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
 Chitnis, Chetan
 Miller, Louis H.
 Peterson, David S.
 Su, Xin-zhaun
 Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/153.273
 FILING DATE: 21-May-2002
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/210.288
 FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
 NAME: Fuller, Michael
 REGISTRATION NUMBER: 36,516
 REFERENCE/DOCKET NUMBER: NIH21.1FMDV1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 921 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium falciparum
 SEQUENCE DESCRIPTION: SEQ ID NO: 8:
 US-10-153-273-8

Query Match 30.5%; Score 521; DB 9; Length 921;
 Best Local Similarity 37.9%; Pred. No. 8.3e-36;
 Matches 127; Conservative 46; Mismatches 106; Indels 56; Gaps 13;

QY 3 DGSNEISGNCNPKESYPDMDC--KKNIDNSHSGACMPRRKOKLCYRDLTGGEIRKPEDILT 61
 DB 407 NGRTTVEGCNCKESYPMDC--KKNNIDSHSGACMPRRKOKLCYH-----EKTMT 457
 QY 62 K-----FINCAKETHFAHMKYKRD--NVNAEN--ELKSGKIPGFRKQMYTTFGDFR 110
 DB 488 NTELKVAIFICAAAEFTELWQNYKRDKNAGNAEDLEKLGILPEDEKRMETFPADYR 517
 QY 111 DIFGSTDISCRYIKDTISOTIKSLGDDATEKEGDTHTD-----NKKLDQEWITIG 162
 DB 518 DICLGTDIS--KKDTSKGV-----GKVKCMLDVFYKISIRYKRSWMEWING 564
 QY 163 PRWGMICALTLNGLS-----ESEKNNIIDYSYNKLNNAEKDDCCLEKFPASRPOFLRW 216
 DB 565 PIVWGMICALSYDPSLNNVNPETHKKLTBEGNNNEFYIPSDSTLSRSEKPPOLRW 624
 QY 217 YVENSDFECPRRKLEKVEDYICAKDYESCKNNKSNNSCVK--CKEYENITGKKT 273
 DB 625 LTEWGENCKBOKK-EYKV--LAKCKDCDDVDGKCKNGKCVACKDCKQYHSMIGIMID 681
 QY 274 QYESQEGKFNTERK--OKKPEYNSYSKRDASEYLK 306
 DB 682 NYKKQKGRYEVKRIPLKEDKDVKNSDDARDYLK 716

RESULT 3

US-09-924-154-14
 Sequence 14, Application US/09924154
 Patent No. US20020127241A1
 GENERAL INFORMATION:

APPLICANT: Natum, David L.
 Applicant: Sim, Kim L.

TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
 FILE REFERENCE: 05213-0465 43170-262105
 CURRENT APPLICATION NUMBER: US/09/924.154
 PRIOR FILING DATE: 2001-08-07
 PRIOR APPLICATION NUMBER: US 60/223,525

PRIOR FILING DATE: 2000-08-07
 NUMBER OF SEQ ID NOS: 17
 SOFTWARE: Patentin version 3.1
 SEQ ID NO 14

LENGTH: 1143
 TYPE: PRT

ORGANISM: Mammalian
 US-09-924-154-14

Query Match 19.2%; Score 327.5; DB 10; Length 1143;
 Best Local Similarity 27.8%; Pred. No. 2.8e-19;
 Matches 85; Conservative 46; Mismatches 118; Indels 57; Gaps 10;

```

OY      10  GCMPE--ESYDPMDSCKKINDSHSGACMPPRKCLCYRD--LTGGEGIKRPEDILTKFN 65
Db      461  GCATKSNMESNMCTGTFTNKKFPGTEPPRRPTLCIGRTYLLHRGHE----EDYEHHLG 516
OY      66  CAAKETHFAWHKKYKRDVNAENELKSGK1PEGFRRKMYTFPGDFRDIFFGTDISSCRYIK 125
Db      517  ASIYEAQLLTKYKKREKDEMLCSI-----IQNSYADLADILKSGDILIDYQK 564
OY      126  DTSGTIRSKIGDQATEKGTHTIDDKKL--QEWWTIHCPKTIWEGMLCALTLGSESEK 183
Db      565  KMEENLKNVKKDKRRNE-----ESLKIIFREKMDENKENWVKWSAVLKN-----K 610
OY      184  NILDDYSYNKLNNAEKDDCKLEKFAKSPQFLRWYVMSDPECRERKKLDEKVEDYICAK 243
Db      611  ETQCDY-----DKFQATIPQFLRWYVMSDPECRERKKLDEKVEDYICAK 243
OY      244  DYSCCKNNKSNNSCVKCKEYENYITGKTKQYESQEGKFNTEKROK--PEYNSYSKDAAS 302
Db      656  ----KDCDCEKCNCKKCEYKRWIDLRKSEYKQVYTKDKDKKKMYDNIDEVKNKEAN 710
OY      303  EYLKDK 308..
Db      711  VYLRK 716

RESULT 4
US-10-153-273-6
Sequence 6, Application US/10153273
Patent No. US2002169305A1
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
            Chitnis, Chetan
            Miller, Louis H.
            Peterson, David S.
            Su, Xin-zhaun
            Wellem, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SPOUNCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbie Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/153,273
FILING DATE: 21-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 749 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

```

```

; HYPOTHETICAL: NO
; ORIGINAL SOURCE: ORGANISM: Plasmodium falciparum
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-153-273-6

Query Match      15.7%; Score 268.5; DB 9; Length 749;
Best Local Similarity 26.8%; Pred. No. 1.7e-14;
Matches 85; Conservative 45; Mismatches 116; Indels 71; Gaps 15

QY      6 NEISGCPKES--YPDWC-KKNIDNSHGACMPPRKLC-----VRDLTGGEIKRP 56
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     109 SSVFGCKTKIKSKVKKKMNCYSNNKTKEGVGPAPRROLCLGYIFLIRDSNEE----- 163
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY    57 EDILTRFLNCAKETHEFWMH-KYKKNVNAENELSGRIPIGFGRKOMYTTFGDFFDIPFG 115
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     164 ---LKDHINKAA--NYEAMHLEKEKENNGDKICIA-----ILGSTADIGDIYRG 208
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   116 TDI---SSCRYIKDTSTQITKSRLGDOATTEKGDIHIDNKKLQEWMTIHGRPIWGMICA 172
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     209 LDVMRDITNTKLSEFKQIFMGGSRRKKQ-----NNNERKNKWEMQRNLIMSSMY-- 260
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   173 LTNGISESEKKNIIODSYNKLNNAEKDCCLEKFAASKPOFLRMVVEMSDFCERPKKLE 232
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     261 -----KHIFGKTCTCRHNPFK-----IPOPRLMLKEWGDEFCBEGITEV 300
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   223 DKVEDVICAKADYECSKNNKSNNSCVKYCKEYENITYGKTQYESQSECFNTEROKRKPE 292
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     301 KOLEKIC---ENKNCSEKK---CKNKCSSYEKMIKERKNMEYNLSQSKFPDCKLTK-K 351
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   293 YNSISK-KDASEYTLKDK 308
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     352 NLLYNKFEDSKAYLRSE 368
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
US-09-924-154-15

Sequence 15, Application US/09924154
Patent No. US20020127241A1
GENERAL INFORMATION:
APPLICANT: Narum, David L.
APPLICANT: Sim, Kim L.
TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
FILE REFERENCE: 05213-0465 43170-262105
CURRENT FILING DATE: 2001-08-07
PRIOR APPLICATION NUMBER: US/09/924,154
CURRENT FILING DATE: 2000-08-07
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15
LENGTH: 1086
TYPE: PRT
ORGANISM: Mammalian
US-09-924-154-15

Query Match      15.7%; Score 268.5; DB 10; Length 1086;
Best Local Similarity 26.8%; Pred. No. 2.7e-14;
Matches 85; Conservative 45; Mismatches 116; Indels 71; Gaps 15

QY      6 NEISGCPKES--YPDWC-KKNIDNSHGACMPPRKLC-----VRDLTGGEIKRP 56
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     351 SSVFGCKTKIKSKVKKKMNCYSNNKTKEGVGPAPRROLCLGYIFLIRDSNEE----- 405
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   57 EDILTRFLNCAKETHEFWMH-KYKKNVNAENELSGRIPIGFGRKOMYTTFGDFFDIPFG 115
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     406 ---LKDHINKAA--NYEAMHLEKEKENNGDKICIA-----ILGSTADIGDIYRG 450
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   116 TDI---SSCRYIKDTSTQITKSRLGDOATTEKGDIHIDNKKLQEWMTIHGRPIWGMICA 172
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     451 LDVMRDITNTKLSEFKQIFMGGSRRKKQ-----NNNERKNKWEMQRNLIMSSMY-- 502
       : : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY   173 LTNGISESEKKNIIODSYNKLNNAEKDCCLEKFAASKPOFLRMVVEMSDFCERPKKLE 232

```

Db 503 -----KHIPKGTCKTGRHNNFEK-----IPQFLRMKEMNDEFCCEENGTEV 542
QY 233 DKVEDYCIKAKOEGCKNNKSNNSCKVCKEYENITGKKTOTYESOEGFNTKROKPE 292
Db 543 KOLEKIC-----EKNKSEKK-----CKNACSYEKWIKERKNEYNLOSKKFDSDKLNK-K 593
QY 293 YNSYSK-KDASEYLKDK 308
Db 594 NMLNKFEDSKAYLRSE 610

RESULT 6
US-10-153-273-21
; Sequence 21, Application US/10153273
; Patent No. US20020169305A1
; GENERAL INFORMATION:
; APPLICANT: Slim, Kim L.
; Chitnis, Chetan
; Miller, Louis H.
; Peterson, David S.
; Su, Xin-zhaun
; Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/153,273
; FILING DATE: 21-May-2002
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,288
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION: }
; NAME: Fuller, Michael
; REGISTRATION NUMBER: 36,516
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-153-273-21

Query Match 15.4%; Score 263.5; DB 9; Length 311;
Best Local Similarity 26.3%; Pred. No. 1.5e-14;
Matches 67; Conservative 19; Mismatches 158; Indels 11; Gaps 4;

QY 22 CKKNIDNSHGACMPRRORIKC---VRDLTQGEIRKPEDILTRKFINCAKETHFAWHKY 78
Db 12 CXXXXXXXXXXACMPRRORIKCLYIIXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 71

QY 79 KKDYNNAENELKSGKIPGFRKOMYTFGDFRDIFFGTDISSCRYIKDTSOTIKSLQDQ 138
Db 72 XXXXXXXXXXXXXXXXXXXXQFLRSMYTFGDFRDIICLNTDIS-----KQNDVXXXXXXXX 126
QY 139 ATTEGSDPHIDNNKILQEWMTIHGPKTWEGMLCAL-TNGLSESEKKNILQDYSTNKLNA 197
Db 127 XXXXXXXXKSPSGLSRQDEWMTNGPELWKGMLCALXXXXXXXXXXXXXXXXXXXXX 186
QY 198 EKDDCCELEKFAKPOFLRMVYEWSDPFCRERKLEDKVEDYCIKAKDYEGCKNNKSNNSC 257
Db 187 XXXXXXXXXXXXXKQFLRMVLEWGEBCAERKKNENIKDACCXXXXXXXXXXCXXK--HRC 244
QY 258 VKVCKEYENITGKK 272
Db 245 NQACRAVOEYENRK 259

RESULT 7
US-09-924-154-16
; Sequence 16, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
; APPLICANT: Narum, David L.
; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
; FILE REFERENCE: 05213-0465 43170-262105
; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US/09/924,154
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Version 3.1
; SEQ ID NO 16
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Mammalian
US-09-924-154-16

Query Match 15.4%; Score 262.5; DB 10; Length 972;
Best Local Similarity 24.7%; Pred. No. 7.5e-14;
Matches 78; Conservative 51; Mismatches 130; Indels 57; Gaps 11;

QY 11 CNPESYPMDCCKKNIDNSHGACMPRRORIKLVRLITOGGELIRKEDILTRFIN-----C 66
Db 138 CKKGNKTIWQCTINEHFKDPPDYCGPFRQDCLGNLDR-DEKKNVND-LKKFLNEITIG 195
QY 67 AAKETHFAWHKYKKNVNAENELKSGKIPGFRKOMYTFGDFRDIFFGTD-----IS 119
Db 196 IRDEGKFLLEKRYK-----NMHENMYLDERACKYLYNSPDYKKNILIKDMMRDNSIK 249
QY 120 SCRYIKDTSOTIKSLQDQATTEKGDTHIDNNKLODEWMTIHGPKIWEGLCALINGLSE 179
Db 250 TENILKGNFEGIKANIVSMYPS-YADLSIDEPFR--HWDQNNKQLEAISEFYKG-- 303
QY 180 SEKKNILODYSYNNKLNNAEKDCELEKFAKPOFLRMVYEWSDPFCRERKLEDKVEDYC 239
Db 304 -----NHTGVLMBDDNDNGLMFMEDKKNDFIDILKNNVDYIKEC 345
QY 240 IAKKDYEGCKNNKSNNS-----CVKCKEYENITGKKTOTYESOEGFNTKROKPE 292
Db 346 IDRK---VKSPPSPSPSDVATVCKNSCTDYDKMIINRKEKYMOSKXY---KRDRSLF 398
QY 293 YNSYSKKAASEYLKDK 308
Db 399 NNVTONIKRPEYLSMK 414

RESULT 8
US-09-924-154-13
; Sequence 13, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:

```

: APPLICANT: Narium, David L.
: APPLICANT: Sim, Kim L.
: TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
: FILE REFERENCE: 05213-0465 43170-262105
: CURRENT APPLICATION NUMBER: US/09/924,154
: CURRENT FILING DATE: 2001-08-07
: PRIOR APPLICATION NUMBER: US 60/223,525
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: Patentin version 3.1
: SEQ ID NO 13
: LENGTH: 1421
: TYPE: PRT
: ORGANISM: Mammalian
: US-09-924-154-13

Query Match      15.3% Score 261; DB 10; Length 1421;
Best Local Similarity 27.2%; Pred. No. 1,6e-13;
Matches 81; Conservative 46; Mismatches 101; Indels 70; Gaps 14;

20 WDCKKNIDNSHSGACMPRRKOKLCVRDLTGGGIRKPED-----ILKFKICAKETHFA 74
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
474 WECKPKPKLSTKYCVPRROELCL-----GNIDRIYDKNLMIKEHILAIAYESRL 527
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
75 WHKXKDNVNAENELKSGKIPGFRKQMYTFGDFRDIFFGTDISCRYIKDTISQITKSK 134
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
528 KRRYK-----NKDDKEVCKIINKEFADRIDIGTD-----YWNDSL--NRK 567
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
135 LGDOATTEKGDTHID-DNKKL--OEWMTIHGPKIWEGMLCALTNGLSESEKKNILQDYSY 191
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
568 LVGGININSNYVHNKQNDKLFDEMKVYIKDVA-----NVISWVEK----- 610
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
192 NKLNAKDDCCLEKFAKPOFLRWYVMSDEFCEKREKKLEKVEDVICAKDEYEGGKNN 251
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
611 -----DKYCKEDIEDINIQFFRFSEMGDDYQDXTKM---IETLKECKE-KPEE-- 658
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
252 KSNNSCVKCEYENYITGKKTQYESQEGFNTKROKKKPEYNSYKDKASE---YLK 306
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
659 --DNCKRKCNSYKEWISKKEEYKQAKQY--OEOYQGNMYKMSFKSIKPEVYLK 712
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

RESULT 9
US-10-153-273-2
: Sequence 2, Application US/10153273
: Patent No. US20020169305A1
: GENERAL INFORMATION:
: APPLICANT: Sim, Kim L.
: Chitnis, Chetan
: Miller, Louis H.
: Peterson, David S.
: Su, Xin-zhaun
: Wellens, Thomas E.
: TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Knobbe Martens Olson & Bear
: STREET: 620 Newport Center Drive 16th Floor
: CITY: Newport Beach
: STATE: California
: COUNTRY: US
: ZIP: 92660
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/153,273
: FILING DATE: 21-MAY-2002
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/09/210,288
```

```

: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Fuller, Michael
: REGISTRATION NUMBER: 36,516
: REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 235-8550
: TELEFAX: (619) 235-0176
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1115 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL SOURCE:
: ORIGINAL SOURCE:
: ORGANISM: Plasmodium vivax
: SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-153-273-2

Query Match      15.2% Score 259; DB 9; Length 1115;
Best Local Similarity 24.9%; Pred. No. 1,8e-13;
Matches 89; Conservative 38; Mismatches 127; Indels 104; Gaps 14;

1 GNDGSNEISGCP-----KESYPMDCCKNIDNS 29
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
223 GNSRKNSSNGSNPFDIDHKTISSAIINHAFLQNTVAKNCYKRRKRRBDMDC-----NT 277
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
30 HSGACMPRRKOKLCVRDLT---OGGELRKPEDIL-----TKFINKAETHEPAW--- 75
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
278 KVDICIPDRRYQCLMKETLNVNNTDINFHDIKFLKRLIYDAVAGDILLKLN 337
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
76 HKYKKNVNAENELKSGKIPGFRKQMYTFGDFRDIFFGTDISCRYIKDTISQITSKL 135
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
338 YRYNKD-----FCKDIRWSLGDGDIIMGTDMGIGSKVVENNLNLSIF 381
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
136 GDOATTEKGDTHIDNKKLOEWMTHGPKIWEGMLCALTNGLSESEKKNILQDYSYKLN 195
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
382 G---TDEKA-----QRRKQWNSKNAQIWTAMYSYKRRL---KGNFIWICKLWAV 428
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
196 NAEKDDCCLEKFAKPOFLRWYVMSDEFCEKREKKLEKVEDVC---IKADYEGCKNNK 252
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
429 NIE-----POLYRWIRMGGRDYSELFTEYOKLEKCDGKINITYDKRVCK--- 473
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
253 SNNSCVKCEYENYITGKKTQYESQEGF-----NTEKROKKPEYNSY--KDKASEY 304
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
474 -VPCQNAKSYDQWITTRKNQMDVLSNKFISVNAEKVQAGIIVTPYDILKQELDEF 530
   |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

RESULT 10
US-09-924-154-17
: Sequence 17, Application US/09924154
: Patent No. US20020127241A1
: GENERAL INFORMATION:
: APPLICANT: Narium, David L.
: APPLICANT: Sim, Kim L.
: TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
: FILE REFERENCE: 05213-0465 43170-262105
: CURRENT APPLICATION NUMBER: US/09/924,154
: CURRENT FILING DATE: 2001-08-07
: PRIOR APPLICATION NUMBER: US 60/223,525
: PRIOR FILING DATE: 2000-08-07
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: Patentin version 3.1
: SEQ ID NO 17
: LENGTH: 1501
: TYPE: PRT
: ORGANISM: Mammalian
: US-09-924-154-17

Query Match      15.2% Score 258.5; DB 10; Length 1501;
Best Local Similarity 23.4%; Pred. No. 2,8e-13;
```

Matches 73; Conservative 61; Mismatches 119; Indels 59; Gaps 13;

QY 11 CNPKESYPMDOCKRIHDS--HSGACMPPRROKLCVRLDTQ--GGEIRKPED--ILTFEI 64
 Db 435 CLEKEFPKCDKNSFETVHHKGVCSPPRQEGCLGNLYLNDLDIYVHNSQLLEII 494
 QY 65 NCAAKETFAHKKYK--DYNVAENELKSGKIPGFRKQMYTGTGDRDIFFGDI---- 118
 Db 495 MASQEGKLLMKKIGTIIIDNMA-----CKYINDSYVDYKDIYTGNDLWMDN 541
 QY 119 SSCRIKDTQTIKSLGDQATTEKGDTHIDNKKLOEMWTIHGPKLMEGMLCALNGLS 178
 Db 542 NSIVQNNLNIIFERNGVKGRNKLRTIKELKNV--WILNKKYVESNRC--GID 595
 QY 179 ESEKKNLQDYSYKLNNAEKDDCCLEKFAKSPQFLRWYVWSDSEFCRERKKLEKVEDV 238
 Db 596 EVDPR-----KTCER---IDELNMPQFFWFSQMAHFFCEKEKEYELKNDK 641
 QY 239 CIRAKDEGCKNNKS--NNSCVKCEYENYITGKKTOYESQEGKNTKROKKPEYNS 295
 Db 642 CTG-----NNGSKLQDQKTCQNVCTNNMYTTRKRLAYELQSVKDKDKL---FSL 690
 QY 296 YSKDASEYLKD 307
 Db 691 AKDKNVTFLKE 702

RESULT 11

US-10-153-273-4
 ; Sequence 4, Application US/10153273
 ; Patent No. US20020169305A1
 ; GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

Chilnis, Chetan
 Miller, Louis H.
 Peterson, David S.
 Su, Xin-zhaun
 Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/153,273
 FILING DATE: 21-May-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/210,288
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Fuller, Michael

REGISTRATION NUMBER: 36,516
 REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1435 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium falciparum

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-10-153-273-4

Query Match 14.7%; Score 251; DB 9; Length 1435;

Best Local Similarity 26.3%; Pred. No. 1,1e-12;

Matches 79; Conservative 47; Mismatches 100; Indels 74; Gaps 14;

QY 20 WDKKKINDNSHSGACMPPRROKLCVRLDTQ--GGEIRKPED--ILTFKINCAAEETHA 74
 Db 474 WECKNYYILSTKVCYCPPRROELC-----GNIDRIYKNTLMITKEHILALAIYESRL 527
 QY 75 WHYKKDYNVAENELKSGKIPGFRKQMYTGTGDRDIFFGDIDISSCRITKTSQTSK 134
 Db 528 KRKYK-----NKDKREVCKIINKTFADIRDIIGTD-----YMNLSN--RKL 568
 QY 135 LGDQATTEKGDTHIDNKK-----LOEMWTIHGPKIWSGMLCALNGLSESEKKNILQDY 189
 Db 569 VKKINTNSK--YVHNRKKNKDLFRDEWKKVIRKIDW-----NVISWVK----- 610
 QY 190 SYNKLNNAEKDDCCLEKFAKSPQFLRWYVWSDSEFCRERKKLEKVEDVCIKADYEGCK 249
 Db 611 -----DKTYCKEDDLEINIPQFFWFSQMAHFFCEKEKEYELKNDK 658
 QY 250 NKSNNSCYKVCKEYENYITGKKTOYESQEGKNTKROKKPEYNSYKSKDASE--YLK 306
 Db 659 ----DNDKSKNSYKWKWISKRKEYNKQAKY--QFYQKGNMYKMYSEFSINPEVYLK 712

RESULT 12

US-10-153-273-10
 ; Sequence 10, Application US/10153273
 ; Patent No. US20020169305A1
 ; GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

Chilnis, Chetan
 Miller, Louis H.
 Peterson, David S.
 Su, Xin-zhaun
 Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/153,273
 FILING DATE: 21-May-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/210,288
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Fuller, Michael

REGISTRATION NUMBER: 36,516
 REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 10:

```

      FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
      NAME: Fuller, Michael
      REGISTRATION NUMBER: 36,516
      REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
TELECOMMUNICATION INFORMATION:
      TELEPHONE: (619) 235-8550
      TELEFAX: (619) 235-0176
      INFORMATION FOR SEQ ID NO: 17:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 324 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
      MOLECULE TYPE: peptide
      HYPOTHEICAL: NO
      ANTI-SENSE: NO
      FRAGMENT TYPE: Internal
      ORIGINAL SOURCE:
      SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-10-153-273-17

Query Match      13.0%; Score 221; DB 9; Length 324;
Best Local Similarity 25.5%; Pred. No. 6,3e-11;
Matches 69; Conservative 17; Mismatches 119; Indels 66; Gaps 8;

QY      33 ACMPRRPKTCVADL--TGGELRKRKEDLTLFINCARETHFAHMKYKKDQVNAENELK 90
      11:|||||:|:|
Db      24 ACIPPRQKCLHYLAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 83
      11:|||||:|:|
QY      91 SGKIPGFRKQMYVTGDFRDIFFGTDISCRVIRKTS-----QTIKSLGDOAT 140
      11:|||||:|:|
Db      84 XXXXXXXFKRQMYETPADYRDLICLGTDIS--KKDTSXXXXXXXXXXXXXXXXXISNIR 140
      11:|||||:|:|
QY      141 TEGDPIHIDNKKLQEWMTIHGKRIWEGMLCAL-----TNGLSEKKNILODYSYKL 194
      11:|||||:|:|
Db      141 YR------SWETNGPVLEWGLCALXXXXXXXXXXXXXXXXXXXXXXXXX 188
      11:|||||:|:|
QY      195 NNAEKDDCLEKFAKPOFLRMVVEWSEDFCERKKLEDKV-----ED 237
      11:|||||:|:|
Db      189 XXXXXXXXXXXXXXXXRPQFLRMVLEWGENFCRQKK-EYKVLACXXXXXXXXXXCCX 247
      11:|||||:|:|
QY      238 VCIAKADYEGCKNNKSNNSCVAKCYEYNT 268
      11:|:|:|
Db      248 XCVACKDQ-----CKQYHSMI 263
      11:|:|:|

RESULT 14
US-10-153-273-15
: Sequence 15, Application US/10153273
: Patent No. US20020169305A1
GENERAL INFORMATION:
APPLICANT: Slim, Kim L.
      Chlunis, Chetan
      Miller, Louis H.
      Peterson, David S.
      Su, Xin-zhaun
      Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
      AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
      ADDRESSEE: Knobbe Martens Olson & Bear
      STREET: 620 Newport Center Drive
      CITY: Newport Beach
      STATE: California
      COUNTRY: US
      ZIP: 92660
COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10-153,273
FILING DATE: 21-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1PWDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-10-153-273-15

```

[illegible]

RESULT 15
 US-10-153-273-13
 : Sequence 13, Application US/10153273
 : Patent No. US20020169305A1
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Sim, Kim L.
 :
 : Chitnis, Chetan
 :
 : Miller, Louis H.
 :
 : Peterson, David S.
 :
 : Su, Xin-zhaun
 :
 : Wellem, Thomas E.
 :
 : TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 : AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 :
 : NUMBER OF SEQUENCES: 37
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Knobbe Martens Olson & Bear
 : STREET: 620 Newport Center Drive 16th Floor
 : CITY: Newport Beach
 : STATE: California
 : COUNTRY: US

```

? ZIP: 92660
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC COMPATIBLE
? SOFTWARE: PatentIn Release #1.0., Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/10/153,273
? FILING DATE: 21-May-2002
? CLASSIFICATION: <Unknown>
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US/09/210,288
? FILING DATE: <Unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Fuller, Michael
? REGISTRATION NUMBER: 36,516
? REFERENCE/DOCKET NUMBER: NIH21.1EMDV1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619) 235-0176
? TELEFAX: (619) 235-0176
? INFORMATION FOR SEQ ID NO: 13:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 291 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? HYPOTHEetical: NO
? ANTI-SENSE: NO
? FRAGMENT TYPE: internal
? ORIGINAL SOURCE:
? SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-10-153-273-13

```

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 14:59:06 ; Search time 24.3035 Seconds

(without alignments)
1218.319 Million cell updates/sec

Title: US-10-087-013-11

Sequence: 1 GNDGSENEISGNCMPKESYDPM.....KKREYNSYSKKDASEYLKDK 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	894	52.4	3006	2	T28625	variant-specific s
2	669.5	39.2	3078	2	T28432	variant-specific s
3	600	35.2	3026	2	T28431	variant surface pr
4	526	30.8	2706	2	T28155	variant-specific s
5	433	25.4	1711	2	C71625	variant-specific s
6	395	23.2	2042	2	T18399	variant-specific s
7	385	22.6	2228	2	T14029	variant-specific s
8	383	22.5	2212	2	T28157	erythrocyte membra
9	367.5	21.5	2182	2	T28634	variant-specific s
10	361.5	21.2	2647	2	T28161	hypothetical prote
11	354	20.8	2135	2	T14602	variant-specific s
12	352	20.6	2664	2	T28626	variant-specific s
13	342.5	20.1	2197	2	B71600	variant-specific s
14	311	18.2	2924	2	T18378	variant-specific s
15	298.5	17.5	1729	2	T18396	erythrocyte membra
16	259	15.2	1070	2	T30848	Duffy receptor - p
17	251	14.7	1435	2	A37793	erythrocyte-blidin
18	246.5	14.4	1153	2	T28652	erythrocyte-blidin
19	227.5	13.3	1045	2	T18373	erythrocyte blidin
20	178	10.4	778	2	A35970	erythrocyte-blidin
21	121	7.1	2523	2	T18477	hypothetical prote
22	117	6.9	1282	2	JE0120	glycoprotein A - m
23	114.5	6.7	508	2	E71620	hypothetical prote
24	113.5	6.5	763	2	G97026	superfamily I DNA
25	111.5	6.5	696	2	G71620	hypothetical prote
26	110.5	6.5	1712	2	C71618	hypothetical prote
27	110.5	6.5	3724	2	T18427	hypothetical prote
28	109	6.4	1939	2	T18372	repeat organellar
29	109	6.4	2391	2	T18410	carbamoyl-phosphat

30	108	6.3	433	2	T25946	hypothetical prote
31	107.5	6.3	219	2	T19897	hypothetical prote
32	107	6.3	1997	2	F71607	DNA helicase II BR
33	106.5	6.2	463	2	S27757	embryonic abundant
34	106.5	6.2	3844	2	T18402	asparagine/asparta
35	106	6.2	1138	2	E71615	probable amine tra
36	105.5	6.2	446	2	T08187	hypothetical prote
37	105	6.2	522	2	C96608	hypothetical prote
38	105	6.2	1302	1	JC6009	surface-located me
39	103	6.0	342	2	T18474	hypothetical prote
40	103	6.0	479	2	T47561	late embryogenesis
41	103	6.0	807	2	T18454	hypothetical prote
42	103	6.0	930	2	D71617	SERA antigen/papai
43	103	6.0	1025	2	S69790	fibronectin-blidin
44	103	6.0	1192	2	A71623	probable secreted
45	103	6.0	1855	2	S41649	DNA polymerase - m

ALIGNMENTS

RESULT 1

T28625
variant-specific surface protein 3 - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T28625
R:Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herrfeldt, J.A.; Peterson, Cell 82, 89-100, 1995
A:Title: The large diverse gene family var encodes proteins involved in cytoadherence
A:Reference number: 220487; MUID:95330813; PMID:7606788
A:Accession: T28625
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3006 <SUX>
A:Cross-references: EMBL:LA0609; NID:9886376; PID:9886377; PIDN:AAA75397.1
C:Genetics:
A:Gene: var-3
A:Introns: 2597/3

Query Match	Score	DB 2;	Length	3006;
Best Local Similarity	52.4%	Pred. No. 4.3e-56;		
Matches 172;	Conservative 38;	Mismatches 91;	Indels 12;	Gaps 4;
QY	5	SNEISGNC-----KESYPMDCCKNIDNSHSGACMPRRQKLCVRDLOGGELRRKEDI	59	
DB	1236	TDDIDGCMOKYKAGKDKYPPGMDNSOIHITHNGACMPRRQKLCVSGLTDRKALEYI	1295	
QY	60	LTFPINCACAKETHPAWHKRYKKNVNAENELKSGKIPGFRKQWYTFGDFRDIFFGTDIS	119	
DB	1296	RTEFIKSALETHPAWHKRYKKNVNAENELKSGKIPGFRKQWYTFGDFRDIFFGTDIS	1355	
QY	120	SCRYIKDTSQ---TIKSLDQDATERGDIHIDNKKLQEWMTIHKRIWEGMICALJNG	176	
DB	1356	THAVISVSPKRVITILEKENDAKYAAKONS---NNEILLDDWMDHGDHIEGMICALTHK	1412	
QY	177	LSESE-KKNILQDYSKKNLNAEKDDCEKFAKPEFLRWYKMSDFEERREKLEDDKY	235	
DB	1413	ISDEKKEKTKNKYSYKRLNESPGSNKVPDFAKPQFLWFLWEGDEFAQREKRAKY	1472	
QY	236	EDVCIRAKDEGCKNNNSNCVCKEYENYINGKTOYESQGRNTEKROKPPYNS	295	
DB	1473	KVSCSDAKDDGCKNNTSNMSVACAVIEDYITKKKVEYTKQKGRDAKIDTKDEYEG	1532	
QY	296	YSKKDASEYLKDK	308	
DB	1533	FSTKDASEYLKKK	1545	
RESULT 2				
T28432		variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)		
N:Alternate names:		erythrocyte membrane binding protein 1 (EMP1)		

RESULT 6
EMBO accession: 118339
Protein: Plasmodium falciparum
Species: Plasmodium falciparum
Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
Accession: 118339
R:RUDLO, J.P.; THOMPSON, J.K.; COWMAN, A.F.
EMBO J. 15, 4069-4077, 1996
Title: The var genes of plasmodium falciparum are located in the subtelomeric region c
Reference number: Z18927; MUID:96324414; PMID:8670911
Accession: 118339
Status: preliminary; translated from GR/EMBL/DBJ
Molecule type: DNA
A:Residues: 1-2042 <RUB>
A:Cross-references: EMBL:U53324; NID:91297090; PID:91297091; PIDN:AA09769.1
Genetics:
A:Map position: 12
Note: 3D7var1

```

Query Match      23.28; Score 395; DB 2; Length 2042;
Best Local Similarity 26.28; Pred. No. 2, 5e-20;
Matches 106; Conservative 65; Mismatches 109; Indels 124; Gaps 15;

OY      24 KNIDNSHG-----ACAPPRRQKLCVADLTQSGEIRKPEDILTFTNCNAKETGFMHHKY 78
          ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       746 RSAGAGSAGSDGACIPIPPRRRLYLHKLI--EGVDITDDSKLRWFLESAVEFFFLMDR 804
          ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

Query Match	22.6%	Score 385:	DB 2:	Length 2228;
Best Local Similarity	24.5%	Pred. NO. 1.2e-19;		
Matches 104; Conservative	59;	Mismatches 123;	Indels 138;	Gaps 14

[illegible]

Db 1276 NCPCGCHGCSYKRWIKRKEIEFKQSNAYGQKTTATRNNGTTFDEKCKTLETWPDAA 1335
 QY 303 EYK 306
 Db 1336 KFLF 1339

RESULT 8

T28157 erythrocyte membrane protein 1 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T28157

R:Yang, S.O.; Wollish, W.S.; Gut, J.; Wu, J.; Ahn, J.; Petersen, C.; Fujioaka, H.; Aikawa, submitted to the EMBL Data Library, July 1995

A:Description: The molecular cloning and DNA sequence analysis of Plasmodium falciparum

A:Reference number: 220479

A:Accession: T28157

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 12212 <YAN>

A:Cross-references: EMBL:U01083; NID:G1517813; PID:G1517814; PIDN:AA06961.1

C:Genetics:

A:Gene: EMP1

Query Match

Best Local Similarity 22.5%; Score 383; DB 2; Length 2212;

Matches 109; Conservative 44; Mismatches 102; Indels 128; Gaps 14;

QY 29 SHSGA-CMPRRQKLCVRDL-----TOGGEI-----RKPEDIL-TKFTN 65

Db 1337 SNSGSLCIPRRRRRLTYQKLDHMASGNTVYVSGAQTPOGTSPPSGKETPSDKLRTAFIQ 1396

QY 66 CAKEHFAWHKYYKKDNVAENELK-----SGKI 94

Db 1397 SAAIEFFFLMDRYKKKEKEKEKKEKANGLVPLNGCPPOPGVTGDSPOSKLQQTGYI 1456

QY 95 PGCFRKYMYTFGDFRDIPEG--TDISSCRYINDTSQTSKLGDAITTEK----- 143

Db 1457 PPPFLRQMYTGLADYADIFFGKNDI-----VIDTKNGDKDIAERKKIKATER 1505

QY 144 ---GDTHIDNKKILOEWMTIHGPKIWEGMLCALT-----NGI-----SES 180

Db 1506 VLKAMDSDPSDEKQNTWEGEHIHMGICATLYKKEDEKGTPLKQNEGLKSLAMDEK 1565

QY 181 EKKNLIDYSYKNNNAEKDDCC-----LEKFAKRPQFLRWYVMSDEFCRERK 229

Db 1566 NKKRPDKQYQDKVLDENSGTSPKTNHVPPTPLNFIISRPYRILEMGEETCRERK 1625

QY 230 KLEDKVEDVCI-----KAK---DYEGCKNNKSN-----SCVAVCKEYENYITGK 271

Db 1626 KRLEKIKVECDDEGKKKQCSGDGCEIRKQDYSTVDFYCPBCGYCRFKRWIKK 1685

QY 272 KTYVESQEGKNTERRKRPPEYN 294

Db 1686 KDEYDKQKQKAVNNKTDARRNN 1708

RESULT 9

T28634 variant-specific surface protein 7 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T28634

R:Su, X.Z.; Heestole, V.M.; Wertheimer, S.P.; Guinet, F.; Herfeldt, J.A.; Peterson, D.S.

Cell 82, 89-100, 1995

A>Title: The large diverse gene family var encodes proteins involved in cytoadherence and

A:Reference number: 220487; MUID:95330813; PMID:7606788

A:Accession: T28634

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-2182 <SOX>

A:Cross-references: EMBL:L42636; NID:G886379; PID:G886380; PIDN:AAA75399.1

C:Genetics:

A:Note: var-7

Query Match

Best Local Similarity 21.5%; Score 367.5; DB 2; Length 2182;

Matches 107; Conservative 60; Mismatches 128; Indels 127; Gaps 16;

QY 10 GCNPKESYDMDG--KKINDNSHG-----ACMPRRQKLCVRDLTQ-----GGE----- 52

Db 874 GPGCEKKEFPNWKCVTPSGVSTATSGDGAICVPPRRRLRYVGLSQMASRGDETFTEVSS 933

QY 53 -----IKRPEDILTKFINKCAKETHEFAWHKYYKKD-----NV 83

Db 934 EATSAPOSSESEKRLATFESAIETFFLHMKYKEKRPATODGAGLGLSPSPGPE 993

QY 84 MNEEL--SGKIPPEGFRKMYTFSDFRDIF--GTDISSCRYIKDTSQTSKLGDAIT 141

Db 994 DPOTLOQTGVIPLPDFLRFMYTLDYKDILYSGNSDTSDTGKQTPSSNDNLKNIVLE 1053

QY 142 EKGDTHIDNK-----KIQEWMTIHGPKIWEGMLCA 172

Db 1054 ASGSTDEQEKKKQIOAKIKITLNGATSGVPPVTKNSYKTPQOTWENIARDINAWCA 1113

QY 173 LT-----NGLS-----ESEKKNILQDYSYNKL-----NNAEKDDC----- 202

Db 1114 LTYKENDARGTSAKIEQNKDKLALMDKANNTPILEKQYINVKLIEDSGAKSNDTIOP 1173

QY 203 CLKFAKSPQFLRWYVMSDEFCRERKLEKVEDVCIK---ANDYEG---CKNNKS-- 253

Db 1174 TLKNFVEIPTFERMLHEWGNCFERAKRLQIKHECDDEDEKQYSGDGEYCEIFSKQ 1233

QY 254 -----NNSCVYKVEYENYITGKKTYVESQEGKNT-----RQKPEYNSKIDAS 302

Db 1234 YNVLDLSSSCAKRRLYKTIIEKKTYERQQAIVEQKSNYENQKQKQTSNNAN 1293

QY 303 EY 304

Db 1294 EF 1295

RESULT 10

T28161 hypothetical protein FC93-varT11-1 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T28161

R:Hernandez-Rivas, R.; Mattei, D.; Sterkers, Y.; Peterson, D.S.; Wellens, T.E.; Scher

Mol. Cell. Biol. 17, 604-611, 1997

A>Title: Expressed var genes are found in Plasmodium falciparum subtelomeric regions.

A:Reference number: 220483; MUID:97154495; PMID:9001213

A:Accession: T28161

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-2647 <HER>

A:Cross-references: EMBL:U07959; NID:G1794255; PID:G1809295; PIDN:AA047438.1

A:Experimental source: strain FCQ27/PMC

C:Genetics:

A:Introns: 2158/3

A:Note: FC93-varT11-1

Query Match

Best Local Similarity 21.2%; Score 361.5; DB 2; Length 2647;

Matches 107; Conservative 44; Mismatches 104; Indels 135; Gaps 14;

QY 29 SHSGA-CMPRRQKLCVRDL-----TOGGEI-----RKPEDIL-TKFTN 65

Db 1338 SNSGSLCIPRRRRRLTYQKLDHMASGNTVYVSGAQTPOGTSPPSGKETPSDKLRTAFIQ 1397

QY 66 CAKEHFAWHKYYKKDNVAENELKSGK-----I 94

Db 1398 SAAIEFFFLMDRYKKKGAIAKKKQKQMDYSLSTADPHNPNVSLVIANPNYKTKCVI 1457

A:Molecule type: DNA
 A:Residues: 1-2197 <GAR>
 A:Cross-references: GB:AE001434; GB:AE001362; NID:93845341; PIDN:AA071996.1; PID:9384534
 A:Experimental source: clone 307
 A:Gene: PFR1055C

Query Match 20.1%; Score 342.5; DB 2; Length 2197;
 Best Local Similarity 24.6%; Pred. No. 1.7e-16;
 Matches 109; Conservative 62; Mismatches 121; Indels 151; Gaps 16;

10 GCNKESEYPMDCCKKINDNS-----HSGA-----CMPPRKQKLCYRDLT-----48
 936 GPGGKEREPNWKCVSSGKSVATAGSSGATGSGDKALCVPPRRRLVYGLTKLTSAG 995
 49 -----OGG-----EIRKPEDILTFINCAKETHPAMHYKK-----80
 996 TSESEPOGSSSESRASVSGNGDDITTTESLRKMFETLTAETEFPLMRYKKEWAOAK 1055
 81 -----DNVNAENEL-KSGKIPEGRKOMYTFGDFRDIFF-----114
 1056 KAEIQRNGLLTGASINLGGDDSNPOTLOKSGTIPDLRLMFTLGDYRDLVRYGA 1115
 115 -----GTDI-----SSGRYIKDTISQTKSLGDAATTEKGD-----THIDNKKLQEW 158
 1116 DKNNGGNITLNSGKNDKQKMEKIOEKI-EQILPTSGNKETRGFQNSVNDROSL---W 1171
 159 TINGPKIMEGMLCALT-----NGL-----SESEKK-----NLL 186
 1172 DRAEHVHGMWCALTYKXDDNGLKGVKKPKQIENPEKLMWNTTKPKPEDEKYOTAKL 1231
 187 QDYSYKLNNAEDDCLEKFSKPOFLRMVYEWSEDFERKRLKEDVEDVCIR-----241
 1232 EDESGERKPDSSASGTAKLDFIKPPYFRYLEEGENCFCKKRTKEMGLIKEQDKNKGRC 1291
 242 -----AKDYEGCKNNKSNNSGVCKYEENYITGKKIYQESQEGKFNTERKOKRP 291
 1292 SGDGKLCNEIIVDKETIFDGLCPTCARHCRTYKMTKNTKRDENFNKSNAYSEOKRYEE 1351
 292 EYNSYSK-----DASEYL 305
 1352 ENDSAKNNKNGVCTLKDAEFL 1374

RESULT 14

T18378

variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)

A:Alternate names: erythrocyte membrane binding protein 1 (EMP1)

A:Species: Plasmodium falciparum

A:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

A:Accession: T18378

A:Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi, T.

Cell 82, 77-87, 1995

A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and

A:Reference number: Z18925; MUID:95330812; PMID:7541722

A:Accession: T18378

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2924 <BAR>

A:Cross-references: EMBL:U27338; NID:9914918; PID:9914919; PIDN:AA060251.1

A:Gene: EMP1

A:Introns: 2476/3

Query Match 18.2%; Score 311; DB 2; Length 2924;
 Best Local Similarity 23.0%; Pred. No. 4.3e-14;
 Matches 105; Conservative 60; Mismatches 120; Indels 172; Gaps 18;

14 KESYPMDCCKKINDNS-----GA-----CMPPRKQKLCYRDL-----47
 860 KTAFTSMKCISSGNNITTESTTKFAGAGTSPGKDTGSCVPPRRRLVYGLHDMAGGET 919
 48 -----TGGG-----IRKPE-DILTFLINCAKETHPAMH 76

Db 920 TEAKSOETSGGQKTPSGNESSPSEKLPQGPTEETKETPESSILHAFVSPRRRLRFLPW 979
 QY 77 KYKKD-----NVNENELKSGKIPEGRKOMYTFGDFRDIFF 113
 Db 980 KFEQWKAQAGAGATGQOTIIGTLDOGGEGETEDKLTGTHITPDLRLQMFYTLGDRDL 1039
 QY 114 FG-TDI-----SSGRYIKDTISQTKSLGDAATTEKGD-----HIDNKKLQEW 158
 Db 1040 VGNLTDIVHTSGNKEDMOIMEAIOKTI-EQILPTSGSSPSPPRTQTOHSEVENPR-KTWW 1097
 QY 159 TINGPKIMEGMLCALTNGL-----SESKKNIIDYISNKLNAEK 199
 1098 NENKKGIMGWCALTYNTDTPSGTAPQIOEVRKIKDENSEKNPKIPQKYDYKLDIT 1157
 200 DDC-----CLEKFSKPOFLRMVYEWSEDFERKRLKEDVEDVCIRAK-D 244
 Db 1158 SDAKTTGSPVPSGKTIPLTDFISRPYFRYLEEGENCFCKKRLKIKIECCGDTG 1217
 QY 245 YEGC-----KNNK-----SNNSCYKVCYEENYITGKKIYQESQEGKFNTERK 287
 Db 1218 HEHCSGDGYDCTRTDADRNDKRVLDNCRDCHIOCKRYKRWIDIKFDEVHKOEKY-----1272
 QY 288 OKPEYNSYSK-----KDASEYLKD 307
 Db 1273 --GEYDKLTKDKSSGDNCKKDIKHKHSAVFLKE 1307

RESULT 15

T18396

erythrocyte membrane protein variant 2 - malaria parasite (Plasmodium falciparum) (fr

A:Species: Plasmodium falciparum

A:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

A:Accession: T18396

A:Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi,

Cell 82, 77-87, 1995

A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen a

A:Reference number: Z18925; MUID:95330812; PMID:7541722

A:Accession: T18396

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1729 <BAR>

A:Cross-references: EMBL:U27339; NID:9914920; PID:9914921; PIDN:AA09134.1

A:Gene: EMP1

A>Note: var-2

Query Match 17.5%; Score 298.5; DB 2; Length 1729;
 Best Local Similarity 23.1%; Pred. No. 1.9e-13;
 Matches 107; Conservative 55; Mismatches 126; Indels 175; Gaps 20;

14 KESYPMDC-----KKNIT-----DNHSGA-CMPPRKQKLCYRDL-----T 48
 Db 860 KTAFTSMKCISSGNNITTESTTKFAGAGTSPGKDTGSCVPPRRRLVYGLHDMAGGETEAKSOET 919
 QY 49 OGGG-----DNVNAENELKSGKIPEGRKOMYTFGDFRDIFFGDTSSGR 80
 Db 920 SCGQKTPSGNESSPSEKLPQGPTEETKETPESSILHAFVSPRRRLRFLPWKFKQWKA 979
 QY 81 -----DNVNAENELKSGKIPEGRKOMYTFGDFRDIFFGDTSSGR 122
 Db 980 QHGAGATGLOLPGVTVDDSDPDPOLOLRGINBNDFLRQMFYTLGDRDICTGCD---R 1035
 123 YI-----KOTISQTKSLGDAATTEKGD-----DHIDNKKLQ 155
 Db 1036 DIVGDTIVSTTEBESITKTKSKITIEGLKQYATSPSPRDTSSRTPVHROTSEKTPQ-Q 1094
 QY 156 EMMTIRGPKIMEGMLCALT-----NGLSESEKK-----NIDQVSYKLNNAE---KDDCC 203
 Db 1095 TMMEANGPHTMNMTCALTFYEDSGAIGQPPQKVEDADVLEKIKRPTANGIKWYLEKEDNT 1154
 QY 204 -----LEKFSKPOFLRMVYEWSEDFERKRLKEDVEDVCIT 240

Db 1155 SSAMPTSSSSSGSNDPINTFPKLTETFEVEIPTFFRYLHEWGONFCERMRRLKQIYKECKV 1214
OY 241 KADY-----EGCKNNKSNN-----CYKVCKEYENYITGKKTO 274
Db 1215 GENGIGRGKOKTPQCCSCYGEDDQLSKYSYDTVADIECPKCAKHCRWYRWIEKKRDE 1274
OY 275 YESOEGKFNTK-----ROKKPEVNSY-----SKKDAEYLK 306
Db 1275 FTEQEKAFPKOKDYVNGNNGKGGDNGFCITLKSLSDAQPLE 1317

Search completed: June 20, 2003, 15:06:23
Job time : 26.3035 secs

THIS PAGE BLANK (USPTO)


```

Db 249 --NLVDTDRH--HSDIIRKSTSRRLIYDVGKDLKTKNNVSEDLCKDIMSIO 304
OY 108 DFDIDFGTDDISSCRITKDTISKLGDOATTEKGTIDHNKLEQWMTIHGPKIME 167
Db 305 DFDIDIMGTDMGIGSLVYQNNLRSLFGTISABE-----DRKK--WMNDHKDKIM 355
OY 168 GMLCALTNGISESEKNNILODYSYKLNNAEKDCCLEKFAK--DQFLRWYVENS-- 221
Db 356 AMI-----LSYKRN-----YSAMN-----CKEDVQIKVEPOIYRMIREGCRDY 396
OY 222 DEFCEKRLKDEKVEVCIAKADYEGCKNNKSNNSGVKCKEVENITGKTOYESQEK 281
Db 397 SEFEREKKLNKCEP-----KLYSTMLICTLPPCNACKSYDEMITSKRWQDVLSTR 451
OY 282 FTEKROKKREYNYSYKSDASEYLKDK 308
Db 452 FSSVKKAKOIKETENIAR--AYDILKQE 476

```

RESULT 2 DR PLAYS

```

ID PVDK_PLAYS STANDARD; PRT; 1070 AA.
AC P22290;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Duffy receptor precursor (Erythrocyte binding protein).
GN PVDR.
OS Plasmodium vivax (strain Salvador I).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID:126793;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE:91187056; PubMed:1849231;
RA Fang X., Kaslow D.C., Adams J.H., Miller L.H.;
RT "Cloning of the Plasmodium vivax Duffy receptor."
RL Mol. Biochem. Parasitol. 44:125-132(1991).
CC - FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP DETERMINANT.
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - SIMILARITY: HIGH, TO P. KNOWLESI DUFFY RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```

```

DR EMBL: M61095; AA:63423.1;
KW Malaria; Receptor; Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 20
FT CHAIN 21 1070
FT TRANSMEM 21 1007-
FT DOMAIN 1008 1025
FT 1026 1070
FT CARBOHYD 183 183
FT CARBOHYD 255 255
FT CARBOHYD 351 351
FT CARBOHYD 420 420
FT CARBOHYD 715 715
FT CARBOHYD 787 787
FT CARBOHYD 825 825
FT CARBOHYD 903 903
FT CARBOHYD 938 938
FT 1070 AA; 119683 MW; CB051DPI3E294603 CRC64;
SQ

```

Query Match 15.2%; Score 259; DB 1; Length 1070;
 Best Local Similarity 24.9%; Pred. No. 3.3e-11;
 Matches 89; Conservative 38; Mismatches 127; Indels 104; Gaps 14;

```

OY 1 GNDGSENLISGNDP-----KESYPMDCCKNNIDS 29
Db 178 GNSRKNSSNGNDYDIDHKKTISALINAFLONTVMKNCNKRKRREDMDIC-----NF 232
OY 30 HGACAPPRKOLCYRDLT---QGEIRKPEDIL-----TKFINCAKETHFAM--- 75
Db 233 KKDVCIPDRRYOLCKMELTNLVNMDTNFHRIDITERKLYLKRKLIYDAVAFGDLKLN 292
OY 76 HKYKKNVNAENELSGKIPSEFRKQMYTFEDDFDIDFGTDDISSCRITKDTISK 135
Db 293 YKXND-----FCKDIRWSLGDFFDIDIMGTDMGIGSLVYQNNLRSLF 336
OY 136 GDOATTEKGTIDHNKLEQWMTIHGPKIMEGMLCALTNGISESEKNNILODYSYKLN 195
Db 337 G---TDEKA-----QRRKQWNNSSKQOITWAMYSYKRL-----KGNFTICLNVA 383
OY 196 NAEKDDCCLEKFAKPOFLRWYVENSDEFCEKRLKEDKVEDVC---IKADYEGCKNNK 252
Db 384 NIE-----POIYRMIREGCRDYVSELPTEYQKLEKCKGKINTYDKKVC--- 428
OY 253 SNNSGVKCKEVENITGKTOYESQEKF---NTEKROKKREYNYS--KDASEY 304
Db 429 -VPCQNAKCKSTQWITRKNQDVLNKNFISVNAEKVQTAGIVTPYDILKQELDER 485

```

RESULT 3

```

ID EBA1_PLAFC STANDARD; PRT; 1435 AA.
AC P19214;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Erythrocyte-binding antigen EBA-175.
OS Plasmodium falciparum (isolate Camp / Malaysia).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID:5835;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE:90377299; PubMed:2204835;
RA Sim B.K.L.;
RT "Sequence conservation of a functional domain of erythrocyte binding antigen 175 in Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 41:293-296(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```

```

DR EMBL: X52524; CA:36756.1;
DR PIR: S11561; S11561.
KW Antigen.
FT DOMAIN 159 1104
FT 1104
FT VARIANT 1031 1031
FT SEQUENCE 1435 AA; 167389 MW; 32A4309021BIC3D6 CRC64;
SQ

```

Query Match 14.7%; Score 251; DB 1; Length 1435;
 Best Local Similarity 26.3%; Pred. No. 1.7e-10;
 Matches 79; Conservative 47; Mismatches 100; Indels 74; Gaps 14;

```

OY 20 WDCKKNIDNSISGACMPRRQKLCVRLTGGELRKPED-----ILTKFINCAKETHFA 74
Db 474 WECKNPYILSTKDVCPVRQELCL-----GNIDRIYDNLMIKELHIAIAYESRIL 527
OY 75 WHRYKKNVNAENELSGKIPSEFRKQMYTFEGDFRIFFGTDDISSCRITKDTISK 134
Db 528 KRRK-----NKDKEVCKTIINKTEADIRDIGTD-----YWNDSN--RKL 568
OY 135 LGDOATTEKGTIDHNK-----LOEWMTIHGPKIMEGMLCALTNGISESEKNNILODYS 189

```

```

Db 569 VGKINTNSK---YVHRKKRNDKLFDEMMKVIKKDW-----NVIWVFK-----610
QY 190 SYKLNNAEKDDCCLEFAKSPQFLRWYVMSDFECREKKLEKVDVYCIKAKDYEGCK 249
Db 611 -----DKTYCKEDDIENIPQFFRWFSMGDDYCODKTKM---JETLVECKE-KPCE 658
QY 250 NKNNSNCVCKEYENITGKTKQYSECKENTEKROKPEYNSYSKKDASE---YLK 306
Db 659 -----DNCKSKCKSYKEMISKKEEYKQAKQY--QYQKGNNTYKMTSEKSIKPEYLYK 712

RESULT 4
PVDA_PLAKN STANDARD: PRT: 1153 AA.
AC P50493;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
Duffy receptor, beta form precursor (Erythrocyte binding protein).
Plasmodium knowlesi.
Eukaryota. Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
NCBI_TaxID=5850;
RN SEQUENCE FROM N.A.
RX MEDLINE=9235776; PubMed=1496004;
RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
RT "A family of erythrocyte binding proteins of malaria parasites.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992).
CC -1- FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP
DETERMINANT.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: HIGH, TO P.VIVAX DUFFY RECEPTOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M90694; AAA29603.1; -
KW Malaria; Receptor; Glycoprotein; Signal; Transmembrane;
FT SIGNAL 1 21
FT CHAIN 22 1153
FT DOMAIN 22 1085
FT TRANSMEM 1086 1106
FT DOMAIN 1107 1153
FT CARBOHYD 134 1153
FT CARBOHYD 134 1153
FT CARBOHYD 179 179
FT CARBOHYD 202 202
FT CARBOHYD 252 252
FT CARBOHYD 348 348
FT CARBOHYD 430 430
FT CARBOHYD 467 467
FT CARBOHYD 576 576
FT CARBOHYD 626 626
FT CARBOHYD 722 722
FT CARBOHYD 847 847
FT CARBOHYD 856 856
FT CARBOHYD 900 910
FT CARBOHYD 910 910
FT CARBOHYD 935 935
FT SEQUENCE 1153 AA; 130471 MW; 6497BD16CE7BEC01 CRC64;

Query Match 14.4%; Score 246.5; DB 1; Length 1153;
Best Local Similarity 25.0%; Pred. No. 2.8e-10;
Matches 81; Conservative 48; Mismatches 116; Indels 79; Gaps 13;

6 NEISGCN--PRESYPMDC--KKNIDNSHGACMPPRRQKLCVRLT---QSGEIRKPED 58
1 : 11 1 1111 11:: 1-1 11 11:::11

```

```

Db 209 NVNRCNNKRRKRGTRDWCPTKKDY-----CIPDRRYLOCNKELTNVNTKTHSHND 261
QY 59 ILTRFINCAKETHFA-----WHYKKDNVNAENELSKGK1PEGRKQMYTTFQGFERD 111
Db 262 ITPLKLNKLEKLYDAVESDILLKTK--NNVISEDLCIKK-----WSLEDFGD 309
QY 112 IFPGTDISCRVYIKDTQSITKSKLGDQATTEKCGTHIDDKKLOEMWTTPKPIWEGMLC 171
Db 310 IINGTDMEGIGISQVYENNNLRITFGTGTQOL-----DKKK--WNNESEKTIWETITL 360
QY 172 ALTNGISESEKKNILQDYSYKLNNAEKDDCCLEFAKSPQFLRWYVMSDFECREKKL 231
Db 361 SVKKKLINGYSAMNCKEDVQIN-----VEPQIYRMIREMGMDYMSLEPKE 404
QY 232 EDKVEDVCIKAKDYEG---CKNKSNNSCVCKEYENITGKTKQYSECKENTEKRO 288
Db 405 QRIKKECKDRKLYLTNLRITGTMSPCND5---CKLYDQWITRRKKQWDVLTSEFSSVKG 460
QY 289 K-----KPEYNSYSK 298
Db 461 QIETENTITAYDILKQELNGFNE 484

RESULT 5
PVDA_PLAKN STANDARD: PRT: 1073 AA.
ID PVDA_PLAKN
AC P22545;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Duffy receptor, alpha form precursor (Erythrocyte binding protein).
OS Plasmodium knowlesi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=5850;
RN SEQUENCE FROM N.A.
RX MEDLINE=9235776; PubMed=1496004;
RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
RT "A family of erythrocyte binding proteins of malaria parasites.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992).
RN [2]
RP SEQUENCE OF 296-1073 FROM N.A.
RX MEDLINE=91004213; PubMed=2170017;
RA Adams J.H., Hudson D.E., Torii M., Ward G.E., Wellens T.E.,
RA Aikawa M., Miller L.H.;
RT "The Duffy receptor family of Plasmodium knowlesi is located within
RT the micronemes of invasive malaria merozoites.";
RL Cell 63:141-153(1990).
CC -1- FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP
DETERMINANT.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: HIGH, TO P.VIVAX DUFFY RECEPTOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M90466; AAA29602.1; -
DR EMBL: M68517; AAA29590.1; -
DR EMBL: M68518; AAA29591.1; -
DR PIR: A35970; A35970.
KW Malaria; Receptor; Glycoprotein; Signal; Transmembrane;
FT SIGNAL 1 21
FT CHAIN 22 1073
FT DOMAIN 22 1007
FT TRANSMEM 1008 1029
FT DOMAIN 1030 1073
FT CARBOHYD 134 134
FT SEQUENCE 1073 AA; 1130471 MW; 6497BD16CE7BEC01 CRC64;

Query Match 14.4%; Score 246.5; DB 1; Length 1153;
Best Local Similarity 25.0%; Pred. No. 2.8e-10;
Matches 81; Conservative 48; Mismatches 116; Indels 79; Gaps 13;

6 NEISGCN--PRESYPMDC--KKNIDNSHGACMPPRRQKLCVRLT---QSGEIRKPED 58
1 : 11 1 1111 11:: 1-1 11 11:::11

```

FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 348 348 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 746 746 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 779 779 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL)
 SQ SEQUENCE 1073 AA; 120683 MW; 3965FC9F46B71808 CRC64;

Query Match 13.3%; Score 227.5; DB 1; Length 1073;
 Best Local Similarity 25.5%; Pred. No. 66-09; Indels 63; Gaps 12;
 Matches 77; Conservative 38; Mismatches 124;

QY 5 SNEISGCP--KESYPMDCCKNNINSHGACMPRRQKLCVRYDT--QGGEIRKPDPI 59
 DB 208 NNVMDCKNDKRRKRGDMDCPAERD-----ICISVRROQLCKGLTNLVNTRTISHNDI 262
 QY 60 LTKFINKCAKETHFAMHK-----YKDNVNAENELKSGKIPESGRKOMYTTGDFPDIFFG 115
 DB 263 TFLKLNLRKRLMYDAVAGDGLLKKNNYQYNKE-----FKDIFMGJGDFDIDIMG 313
 QY 116 TDISSCRYIKDTSQITKSKLGDQATEKGDTHIDNKKLOEWMTIHGKIMEGLCALTN 175
 DB 314 TMEIGTISQYVENNLROVFG--TDERAK---QDRK---QWNSKEHIVRAMFISIRS 364
 QY 176 GLSESEKKNILQDYSYKLNNAEKDCLEKFAKPOFLRMVEMSDCECRKKLEDKV 235
 DB 365 RLKE-----KFWICKKDYVLK---VEPQYRRIKREGRMYSKLKEQSKL 408
 QY 236 EDVCIKADYEGCKNNKNSNS-----CVKCKEYENYITGKKTQYESQEGKENTEKRO 288
 DB 409 NECKASKLY-----NNMAICMLPLCHDACKSYDQMITRKKKQMDVLSTKSSVAKT 460
 QY 289 KK 290
 DB 461 OK 462

RESULT 6
 UBP8_HUMAN STANDARD; PRT; 1118 AA.
 AC P40818;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ubiquitin carboxyl-terminal hydrolase 8 (EC 3.1.2.15) (Ubiquitin thiolesterase 8) (Ubiquitin-specific processing protease 8) (Deubiquitinating enzyme 8).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NX NCBI_TaxID=9606;
 RX [1]
 RP TISSUE=Bone marrow;
 MEDLINE=96051398; PubMed=7584044;
 RA Seki N., Kawarayashi Y., Ishikawa K.-I., Tabata S.,
 RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
 RT "Prediction of the coding sequences of unidentified human genes. II. The coding sequences of 40 new genes (K1AA0041-K1AA0080) deduced by analysis of cDNA clones from human cell line KG-1.";
 RT DNA Res. 1:223-229(1994).
 RL -i- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O = ubiquitin + a thiol.
 CC -i- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: D29956; BAA06225.1; -
 CC MEROPS: C19.011; -
 CC Genew: HGNC:12631; USP8.
 DR MIM: 603158; -
 DR InterPro: IPR001763; Rhodanese-like.
 DR InterPro: IPR001394; UCH-2.
 DR Pfam: PF00442; UCH-1; 1.
 DR Pfam: PF00443; UCH-2; 1.
 DR Pfam: PF00581; Rhodanese; 1.
 DR SMART: SM00450; RHOD; 1.
 DR PROSITE: PS00972; UCH_2_1; 1.
 DR PROSITE: PS00973; UCH_2_2; 1.
 DR PROSITE: PS0235; UCH_2_3; 1.
 DR Ubl conjugation pathway; Hydrolyase; Thiol protease; Multigene family.
 FT ACT_SITE 786 786 BY SIMILARITY.
 FT ACT_SITE 1059 1059 BY SIMILARITY.
 FT ACT_SITE 1067 1067 BY SIMILARITY.
 SQ SEQUENCE 1118 AA; 127523 MW; 8B884B7A842F9A9A CRC64;

Query Match 6.4%; Score 109; DB 1; Length 1118;
 Best Local Similarity 21.6%; Pred. No. 1.9;
 Matches 66; Conservative 34; Mismatches 121; Indels 84; Gaps 14;

QY 60 LTKFINKCAKETHFAMHK-----KDNVNAENEL-KSKTIDEGF-----R 99
 DB 788 MNSILOGLCNAPLADFNRCYODDINRNLGHKCEVAEERGIKALMTQGYRISP 847
 QY 100 KQMYTFGDFPDIFGTDISSCRYIKDTSQITKSKLGD-----QATTEGDT 147
 DB 848 KDEITIGKINDQFAG-----YSQDSQELLFLMDGHELDNKNADNRRYKEENND 900
 QY 148 IDNKKLOEWMTIHGPIWEGMLCALTNGLSES-----EKNILQDYSYKLNNAEK 199
 DB 901 LDDFKAEHAWOKH-KOLNESIIYALFOGQFKSTVCLTGHKRSRFEAFMYLSLPLAST 959
 QY 200 DDC-----CLEKFAKPOFL---RMYVMSDFECRERKKLEDKVD-----VCIKARD 244
 DB 960 SKCTLDQCLRLFSEKELTNNRNYC---SHCRARDSLKTEIKLPLVLLVHLKRS 1015
 QY 245 YEGCKNNKNSNCVKCKEYEN-----YINGKKTQYESQEGKENTEKROKRPENYSYK 299
 DB 1016 YDGRWKKQKLTST---VDPLPNDLSQYVIGPRN-----NLKKNYLFVSNNHGYL 1063
 QY 300 DASY 304
 DB 1064 DGGHY 1068

RESULT 7
 ATRX_HUMAN STANDARD; PRT; 2492 AA.
 AC P46100; P51068; Q15886; Q9NTS3; Q9H021;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Transcriptional regulator ATRX (X-linked helicase II) (X-linked nuclear protein) (XNP) (Znf-HX).
 GN ATRX OR RAD54L OR XR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NX NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5), VARIANT S-1860, AND VARIANTS ATR-X.
 RX MEDLINE=97123494; PubMed=8968741;
 RA Picklets D.J., Higgs D.R., Bachoo S., Blake D.J., Quarrell O.W.J.,
 RA Gibbons R.J.;
 RT "ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome.";

Hum. Mol. Genet. 5:1899-1907(1996).
 [12]
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).
 RX MEDLINE-97386582; PubMed-9244431;
 RA Villard L., Lossi A.-M., Cardoso C., Proud V., Chiaroni P.,
 RA Colliaux L., Schwartz C., Fontes M.;
 RT "Determination of the genomic structure of the XNP/ATRX gene encoding
 RT a potential zinc finger helicase.";
 RL Genomics 43:149-155(1997).
 [13]
 RP SEQUENCE OF 860-2492 FROM N.A.
 RX MEDLINE-95179111; PubMed-7874112;
 RA Stayton C.L., Dabovic B., Gulisano M., Geetz J., Broccoli V.,
 RA Giovannazzi S., Bossolasco M., Monaco L., Rastan S., Boncinelli E.,
 RA Bianchi M.E., Consalez G.G.;
 RT "Cloning and characterization of a new human Xq13 gene, encoding a
 RT putative helicase.";
 RL Hum. Mol. Genet. 3:1957-1964(1994).
 [14]
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
 MEDLINE-94214473; PubMed-8162050;
 Geetz J., Pollard H., Consalez G., Villard L., Stayton C.L.,
 RA Millasseau P., Khrestchatsky M., Fontes M.;
 RT "Cloning and expression of the murine homologue of a putative human
 RT X-linked nuclear protein gene closely linked to PK1 in Xq13.3.";
 RL Hum. Mol. Genet. 3:39-44(1994).
 [15]
 RP SEQUENCE OF 2401-2492 FROM N.A. AND VARIANTS ATR-X.
 RX MEDLINE-95211835; PubMed-7697714;
 RA Gibbons R.J., Picketts D.J., Villard L., Higgs D.R.;
 RT "Mutations in a putative global transcriptional regulator cause X-
 RT linked mental retardation with alpha-thalassemia (ATR-X syndrome).";
 RL Cell 80:837-845(1995).
 [16]
 RP SEQUENCE OF 1375-2492 FROM N.A.
 RA Pearce A., Chapman J.;
 RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 [17]
 RP EZH2 BINDING.
 RX MEDLINE-98167853; PubMed-9499421;
 RA Cardoso C., Tinsit S., Villard L., Khrestchatsky M., Fontes M.,
 RA Colliaux L.;
 RT "Specific interaction between the XNP/ATR-X gene product and the SET
 RT domain of the human EZH2 protein.";
 RL Hum. Mol. Genet. 7:679-684(1998).
 [18]
 RP SUBCELLULAR LOCATION, AND ASSOCIATION WITH PERICENTROMERIC
 HETEROCHROMATIN.
 MEDLINE-20040663; PubMed-10570185;
 McDowell T.L., Gibbons R.J., Sutherland H., O'Rourke D.M.,
 RA Bickmore W.A., Pombo A., Turley H., Gatter K., Picketts D.J.,
 RA Buckle V.J., Chapman L., Rhodes D., Higgs D.R.;
 RT "Localization of a putative transcriptional regulator (ATRX) at
 RT pericentromeric heterochromatin and the short arms of acrocentric
 RT chromosomes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:13983-13988(1999).
 [19]
 RP DISEASE.
 RX MEDLINE-20213147; PubMed-10751095;
 RA Villard L., Fontes M., Ades L.C., Geetz J.;
 RT "Identification of a mutation in the XNP/ATR-X gene in a family
 RT reported as Smith-Rimman-Meyers syndrome.";
 RL Am. J. Med. Genet. 91:83-85(2000).
 [10]
 RP VARIANT ATR-X SER-1713.
 RX MEDLINE-97196774; PubMed-9043863;
 RA Villard L., Lacombe D., Fontes M.;
 RT "A point mutation in the XNP gene, associated with an ATR-X phenotype
 RT without alpha-thalassemia.";
 RL Eur. J. Hum. Genet. 4:316-320(1996).
 [11]
 RP VARIANT JM GLN-2131.
 MEDLINE-96224392; PubMed-8630485;
 Villard L., Geetz J., Mattei J.-F., Fontes M., Saugier-Verber P.,
 RA Munnich A., Lyonnet S.;
 RT "XNP mutation in a large family with Juberger-Warsidi syndrome.";
 RL Nat. Genet. 12:359-360(1996).
 [12]
 RP VARIANTS ATR-X.
 RX MEDLINE-97467722; PubMed-9326931;
 RA Gibbons R.J., Bachoo S., Picketts D.J., Aftimos S., Asenbauer B.,
 RA Bergoffen J., Berry S.A., Dahl N., Fryer A., Keppeler K., Kurosawa K.,
 RA Levin M.L., Masuno M., Neri G., Pierpont M.E., Stanley S.F.,
 RA Higgs D.R.;
 RT "Mutations in transcriptional regulator ATRX establish the functional
 RT significance of a PHD-like domain.";
 RL Nat. Genet. 17:146-148(1997).
 [13]
 RP VARIANT ATR-X LEU-246.
 RX MEDLINE-20123062; PubMed-10660327;
 RA Flichera M., Romano C., Castiglia L., Falla P., Ruberto C., Amata S.,
 RA Greco D., Cardoso C., Fontes M., Ragusa A.;
 RT "New mutations in XNP/ATR-X gene: a further contribution to
 RT genotype/phenotype relationship in ATR/X syndrome.";
 RL Hum. Mutat. 12:214-214(1998).
 [14]
 RP VARIANT SHS LYS-1742.
 RX MEDLINE-99347960; PubMed-10417298;
 RA Lossi A.-M., Millan J.M., Villard L., Orellana C., Cardoso C.,
 RA Prieto F., Fontes M., Martinez F.;
 RT "Mutation of the XNP/ATR-X gene in a family with severe mental
 RT retardation, spastic paraplegia and skewed pattern of X inactivation:
 RT demonstration that the mutation is involved in the inactivation
 RT bias.";
 RL Am. J. Hum. Genet. 65:558-562(1999).
 [15]
 RP VARIANT CMS THR-2050.
 RX MEDLINE-99326061; PubMed-10398237;
 RA Abidi F., Schwartz C.E., Carpenter N.J., Villard L., Fontes M.,
 RA Curtis M.;
 RT "Carpenter-Waziri syndrome results from a mutation in XNP.";
 RL Am. J. Med. Genet. 85:249-251(1999).
 [16]
 RP VARIANTS ATR-X E-175; 179-V-K-198 DEL; S-190; P-219; L-246 AND C-249.
 RX MEDLINE-99219535; PubMed-10204841;
 RA Villard L., Bonito M.-C., Abidi F., Ragusa A., Beloungue J.,
 RA Lossi A.-M., Seaver L., Bonnefont J.-P., Romano C., Flichera M.,
 RA Lacombe D., Hanauer A., Philip N., Schwartz C.E., Fontes M.;
 RT "Evaluation of a mutation screening strategy for sporadic cases of
 RT ATR-X syndrome.";
 RL J. Med. Genet. 36:183-186(1999).
 [17]
 RP VARIANTS ATR-X S-179; L-190; I-194; C-246; F-1552; S-1645 AND C-1847.
 RX MEDLINE-20451413; PubMed-10995512;
 RA Wada T., Kubota T., Fukushima Y., Saitoh S.;
 RT "Molecular genetic study of Japanese patients with X-linked
 RT alpha-thalassemia/mental retardation syndrome (ATR-X).";
 RL Am. J. Med. Genet. 94:242-248(2000).
 [18]
 RP FUNCTION: COULD BE A GLOBAL TRANSCRIPTIONAL REGULATOR. MODIFIES
 CC GENE EXPRESSION BY AFFECTING CHROMATIN. MAY BE INVOLVED IN
 CC BRAIN DEVELOPMENT AND FACIAL MORPHOGENESIS.
 CC -1- SUBUNIT: PROBABLY BINDS EZH2. BINDS ANNEKIN V IN A CALCIUM AND
 CC PHOSPHATIDYLCHOLINE/PHOSPHATIDYLSELINE-DEPENDENT MANNER (BY
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH PERICENTROMERIC
 CC HETEROCHROMATIN DURING INTERPHASE AND MITOSIS, PROBABLY BY
 CC INTERACTING WITH HP1.
 CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS: 1, 2, 3, 4 (SHOWN HERE) AND 5;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: UBICUITOUS.
 CC -1- DISEASE: DEFECTS IN ATRX ARE THE CAUSE OF X-LINKED ALPHA-
 CC THALASSEMIA/MENTAL RETARDATION SYNDROME (ALSO KNOWN AS ATR-X
 CC SYNDROME). ATR-X IS AN X-LINKED DISORDER COMPRISING SEVERE
 CC PSYCHOMOTOR RETARDATION, FACIAL DYSMORPHISM, UROGENITAL
 CC ABNORMALITIES, AND ALPHA-THALASSEMIA. AN ESSENTIAL PHENOTYPIC
 CC TRAIT ARE HEMOGLOBIN H ERYTHROCYTE INCLUSIONS.

```

CC -1- DISEASE: DEFECTS IN ATRX ARE THE CAUSE OF CARPENTER-WIZIRI
CC SYNDROME (CWS), AN X-LINKED, RECESSIVE CONDITION CHARACTERIZED BY
CC MODERATE MENTAL RETARDATION, SHORT STATURE, BRACHYACTYL WITH
CC EXCESSIVE SKIN CREASES, AND WIDENING OF THE KNUCKLES.
CC -1- DISEASE: DEFECTS IN ATRX ARE THE CAUSE OF JUBER-MARSIDI SYNDROME
CC (JM). JM IS A RARE X-LINKED RECESSIVE DISEASE CHARACTERIZED BY
CC (SEVERE) MENTAL RETARDATION, GROWTH FAILURE, SENSORINEURAL DEAFNESS,
CC MICROGENTALISM AND EARLY DEATH.
CC -1- DISEASE: DEFECTS IN ATRX ARE A CAUSE OF SMITH-FITENMAN-MYERS
CC SYNDROME (SPM). CLINICAL FEATURES INCLUDE SEVERE MENTAL
CC RETARDATION, MICROCEPHALY, GROWTH FAILURE, FACIAL ANOMALIES AND
CC BILATERAL CRYPTORCHIDISM. DUE TO THE CLINICAL OVERLAP WITH ATR-X
CC SYNDROME, SOME PATIENTS ORIGINALLY DIAGNOSED AS HAVING SPM, MIGHT
CC BE AFFECTED BY A VARIANT OF ATR-X SYNDROME WHICH LACK HEMOGLOBIN H
CC INCLUSIONS.
CC -1- DISEASE: DEFECTS IN ATRX ARE THE CAUSE OF SUTHERLAND-HAAN X-LINKED
CC MENTAL RETARDATION SYNDROME (SHS). IT IS CHARACTERIZED BY SEVERE
CC MENTAL RETARDATION WITH SPASTIC PARAPLEGIA, MICROCEPHALY, SHORT
CC STATURE AND CRYPTORCHIDISM.
CC -1- SIMILARITY: BELONGS TO THE SMF2/RAD54 HELICASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PHD-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U72937; AAB49970.2; -
DR EMBL: U72938; AAB49971.2; -
DR EMBL: U72935; AAB40698.1; -
DR EMBL: U72904; AAB40698.1; JOINED.
DR -----
Query Match 6.1%; Score 104; DB 1; Length 2492;
Best Local Similarity 18.7%; Pred. No. 11;
Matches 81; Conservative 54; Mismatches 142; Indels 156; Gaps 14;
CC -----
CC 12 NPKESEYPMDC-----KKNINSHSGACMPRRKOLCVRLDTGGELRKP 56
CC 671 NPVYNSDSECENEYKEREKOLSVPRKKDKRNSDSDALDNKPKNLKPSKOSSEYVONSND 730
CC -----
CC 57 EDILKFTINCAKETHFAMHKYKKDNVAEN-----ELKSGKIPEGRKOMYYTFGDFR 110
CC 731 SDEMALILKGVSRMSSSSDPTDINEHTNKTLYDLTKQAGKDKGRKRKSSGSDF 790
CC -----
CC 111 DIFPG-----TDISCRYIKDPSQITK-SKLGDAATPK----- 143
CC 791 DTKGKSAKSSIISKRRQTOSESSNVDSLEKEIKSNKGTGAARTIKRIPNTRKDPSS 850
CC -----
CC 144 -----GDTHT-----DNKKLQEWMTIHG-----PK 164
CC 851 EDEKSKKGMNONGHKNLKTQSESSDAERKQBEFTSSAGTYADKDTIMELRDRPK 910
CC -----
CC 165 IWEGMLCALJNCLSE-----SEKNIILODYSYNKLNAEAKDDCLE 205
CC 911 --KQOASASTGVDLISGKEQSFSLAEYKVAETKEKSHKTKTKCKKVOGLSD--TAE 966
CC -----
CC 206 KFAKSPQFLRWYEVKSDCECRERKLEDKVEDVCIKADYGCCKNNKSNNSCVKCKEYE 265
CC 967 KFLAKDQ-----SDTSEDDKQSKKKGTEKKKPSDFK-----KVIYKEQOYE 1010
CC -----
CC 266 NYITGKK-----TOYEOGFRNTEKROKKEEYNSYSKK----- 299
CC -----
CC 1011 SSSDTEKLPREBEICHPFKGIKQIKNGITDCEKSKKIRPKTKSKKDELSDYAKSKGK 1070
CC -----
CC 300 ---DASEYLKDK 308
CC -----
CC 1071 GDCSDSSEDKSKK 1083
CC -----

```

```

YD33_SCHPO
ID YD33_SCHPO STANDARD: PRT: 278 AA.
AC Q10267;
DT 01-OCT-1996 (Rel. 34; Created)
DT 01-OCT-1996 (Rel. 34; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Hypothetical protein C13G7.03 in chromosome I.
GN SPAC13G7.03
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
XP [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Godle A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkheart G., Aert R., Robben J., Grymonprez B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Meller Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Bierzyn K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Beger P., Zimmermann W., Medler H., Wambut R., Purrelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motter S.,
RA Gilbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Talhada V.A., Garçon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Garattini L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -----
CC -1- SIMILARITY: SOME, TO HISTONES H1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z69729; CAA93591.1; -
DR InterPro: IPR005120; Smg4-UPF3.
DR Pfam: PF03467; Smg4-UPF3; 1.
KM Hypothetical protein.
SQ SEQUENCE 278 AA; 30935 MW; F3BCC571C29C85B CRC64;
CC -----
Query Match 5.9%; Score 100; DB 1; Length 278;
Best Local Similarity 22.7%; Pred. No. 18;
Matches 64; Conservative 33; Mismatches 103; Indels 82; Gaps 13;
CC -----
CC 56 PEDILKFTINCAKETHFAMHKYKKD--NVNAENELKS-----GKIPEGRKOMYY 104
CC 24 PEQYFLQINSFL--PVEVHREFSKGATVGTSELSFAYLKFGOSATVAOEFRRVYQGH 81
CC -----
CC 105 TFGDFRDIFFGDTDISRCRYIKDPSQITKSLGQATTEKEDPHIIDNNKLOEWMTHGPK 164
CC 82 TFDKNNNTYRAVITLAVYOKIPPSKYK-----DSLGLSLDDPKFOE-----K 127
CC -----
CC 165 IWEGMLCALJNCLSESEKKNIILODYSYNKLNN-----AKRDCCLERKFAKSPQFLRW 216
CC -----

```



```

Db 128 VOR-----ESYSQTSASNDVIEKLTQTPPLQYLAELKNNAVENKSKRP----- 177
Oy 217 YVENSDEGCRKRLKLEDEVCYCKADYBECCRNKNSNGVCYKCEKENTITTKTOYE 276
Db 172 ----SKSVKRAKRL-----RLAEKPASNN--KAGKSSQESKSSKAPAE 211
Oy 277 S-----OEGKFNTEKRQCK-----PEXNSYSKPKASEYLKCK 308
Db 212 SAAVIEDKYSDRKSKRKKRKKTPVNS-TASQASENASDK 252

RESULT 9
VDP_MOUSE
ID VDP_MOUSE STANDARD: PRT: 941 AA.
AC Q92120;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE General vesicular transport factor p15 (Transcytosis associated
ne protein) (TAP) (Vesicle docking protein) (Fragment).
VDP.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=129; TISSUE=Thymus;
Han S.J.;
"Identification of mouse TAP (transcytosis associated protein/p15).";
Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR
CC INTERCYTOSOLIC TRANSPORT IN THE GOLGI STACK; IT IS REQUIRED FOR
CC TRANSCYTIC FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO
CC THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY
CC INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR
CC AND TARGET MEMBRANES IN PROXIMITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES
CC BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE (BY
CC SIMILARITY).
CC -!- DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COLLED-
CC COIL) AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.
CC -!- PM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER;
CC PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS.
CC DEPHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLGI MEMBRANE;
CC PHOSPHORYLATION PROMOTES DISSOCIATION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE VDP/USO1/YB1047C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF096868; AAC72967.1; -.
CC MGD; MGI:1929095; VDP.
CC InterPro: IPR000225; Armaad11lo.
CC InterPro: IPR000219; RhogEF.
CC SMART; SM00325; RhogEF; 1.
CC PROSITE; PS50176; ARM_REPEAT; UNKNOWN_1.
CC Transport; Protein transport; Golgi stack; Membrane; Colled coil;
CC Phosphorylation.
CC FT NON_TER 1
CC FT DOMAIN <1 619 GLOBULAR HEAD.
CC FT DOMAIN 620 912 COILED COIL (POTENTIAL).
CC FT DOMAIN 917 941 ASP/GLU-RICH (ACIDIC).
CC MOD_RES 922 922 PHOSPHORYLATION (BY SIMILARITY).
CC SEQUENCE 941 AA; 105152 MW; AB5C10895CD7E508 CRC64;

Query Match 5.8%; Score 99.5; DB 1; Length 941;
Best Local Similarity 19.6%; Pred.No. 7.7;

```

```

Matches      52:  Conservative      32:  Mismatches      92:  Indels      89:  Gaps      6:

OY      43  CVRBLDTGGGELRKREDIL-----TFINCAKETFFAAHKKKKDDVNAENELSKSPIEGF  98
Db      460  CTNLISQGSKSIQTRVGLMLMCLTWSLNSCPAVTFHFLHN-----SANVPFLTGQIAE--  510
OY      99  RKQVYVFGDFPRDIFFGFDISCHYIKDTSQTSIKSGDQDATTEKGTHTDDNKKILQEMW  158
Db      511  -----NIGEEQLVQYG-----  521
OY      159  TTHGPKIWEGMICALTNGLSESEKKNITLQDYSYKNLNAEKDDCLERFASKPOFLRMVY  218
Db      522  -----LCALLGISITVFNDNSLENTKREKLQGLIKRTGKGNFLIEKLGIFSKHE  570
OY      219  EWSE-----FCRRKKLEKDVEDYCIKADYBEGCKNNKSNNSCVKYCKE  263
Db      571  LYSRASQKPPQPFPSPEYMFIDHFETFKLVKLELGVITKA-LYKSSSEDKKEEVEYKRTLEQ  629
OY      264  YENYITGCKKIQYESQEGKFNTERKKO  288
Db      630  HDNIVTHYKNMIREODILOELEKKQO  654

RESULT 10
XY02_METJA
ID      XY02_METJA      STANDARD:      PRT:  1181 AA.
AC      Q60301;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Hypothetical protein MJEC502.
GN      MJEC502.
OS      Methanococcus jannaschii.
OC      Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC      Methanocaldococcaceae; Methanocaldococcus.
OX      NCBI_TaxID:2190;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX      MEDLINE=9633799; PubMed=8688087;
RA      Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA      Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA      Keilavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA      Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA      Scott J.L., Georgann N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA      Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.M., Hanna M.C.,
RA      Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA      Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT      "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT      jannaschii."
RL      Science 273:1058-1073(1996).
CC      -I- SIMILARITY: WEAK IN THE C-TERMINUS, TO ADENINE-SPECIFIC
CC      METHYLTRANSFERASES.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on way
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L77119; AAC37060.1; -.
DR      TIGR: MJEC502; -.
DR      InterPro: IPR002296; N12N6_mtfase.
DR      InterPro: IPR002052; N6_mfase.
DR      PRINTS: PR00507; N12N6MTFASE.
DR      PROSITE: PS00092; N6_MTFASE.1.
KM      Hypothetical protein: Transmembrane; Complete proteome.
FT      TRANSMEM 837 857  POTENTIAL.
SEQUENCE 1181 AA; 138438 MW; 9F961D8F8C6AABDD CRC64;

Query Match      5.8%; Score 99; DB 1; Length 1181;

```

Best Local Similarity 21.5%; Pred. No. 11;
Matches 63; Conservative 41; Mismatches 95; Indels 94; Gaps 13;

QY 41 KICVBDLQGGGIRRPEDILKE-----INCAKETHFAHMY----KKNVNAEN 87
DB 8 KLC-----SDIEIPKEVYNNKMDFKILVGNIIISIQIPELEKFEKLLKGRYNNKS 60
QY 88 ELKSGKIPGCF-RKMVYTFGFRDIFPGTDSGCRVYIKDTSQTKSLGD----- 137
DB 61 DDFGQDLEPFTKRVIVIEPIEFLEFTESEISK-----KSLPGDKKIDPVYNS 109
QY 138 -----QATTEGDTTHIDNNK-----LQEMWTIHGPKTWEGMLCALTNGLS----- 178
DB 110 VENKEILIEAPLEGLSDLNKKDSCGHQVKEMLIKSGVDTG---ATNGLEAVLLHYDDT 166
QY 179 -----ESEKKNILQDYSYNNKLNNAEKDDCCLEK-----FASKPOFLRMVY----- 218
DB 167 IKEIRTLKLNKLSIFEYVLENNKDKD---LENELKOVSEFYCYFSKYEIEYEVATK 223

219 -----EWSDEFCRERKKLEDDVEDV-CIRAKDEGCGKNNKSNVCVCKE 263
224 NIKHKKEITNEFYKEFYKLVGFEDVDKVKKKDSSSEKDKGTAKCLANCIE 276

RESULT 11
NUF1_YEAST
ID NUF1_YEAST STANDARD: PRT: 944 AA.
AC P32380:
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE NUF1 protein (spindle body protein spacer protein SPC110).
GN NUF1 OR SPC110 OR YDR356W OR D9476.3.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE=92176232; PubMed=1541631;
RA Mirzayan C., Copeland C.S., Snyder M.;
RT "The NUF1 gene encodes an essential coiled-coil related protein that
is a potential component of the yeast nucleoskeleton";
RL J. Cell Biol. 116:1319-1332(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=94064779; PubMed=7503995;
RA Kilmarlin J.V., Dyos S.L., Keirshaw D., Finch J.T.;
RT "A spacer protein in the Saccharomyces cerevisiae spindle body
whose transcript is cell cycle-regulated.";
RL J. Cell Biol. 123:1175-1184(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Ravello A., Fulton L., Gattung S., Greco T., Kirsten J.,
RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA Marsden E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA Riffen L., Riles L., Taich A., Trevisan E., Vignati D.,
RA Wilcox L., Woldman P., Vaudin M., Wilson R., Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEOSKELETON. POTENTIAL
IS ESSENTIAL FOR GROWTH.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; TIGHTLY ASSOCIATED WITH THE
NUCLEUS. IT IS PRESENT IN A GRANULAR PATTERN THAT EXCLUDES THE
NUCLEOLUS.
CC -1- PTM: MAY BE REGULATED BY PHOSPHORYLATION EVENTS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: 211582; CAAT7668.1; -
DR EMBL: X73297; CAAS1733.1; -
DR EMBL: U28372; AAB64791.1; -
DR PIR: S26710; S26710.
DR PIR: S34288; S34288.
DR SCD: S0002764; NUF1.
KW Coiled coil; Nuclear protein; Phosphorylation.
FT DOMAIN 164 791
FT DOMAIN 54 59
FT DOMAIN 726 731
FT DOMAIN 742 747
FT DOMAIN 731 944
SQ SEQUENCE 944 AA; 111781 MW; 04FAA074BB8A0BC8 CRC64;

Query Match 5.7%; Score 97.5; DB 1; Length 944;
Best Local Similarity 23.8%; Pred. No. 11;
Matches 81; Conservative 53; Mismatches 127; Indels 79; Gaps 20;

QY 21 DCKRNIDS-----HSGACMPPRQ-----KLC-----YRDL-----TQGEIRK 55
DB 208 DAKKKIVSRKVDHSG-CIEERQMERKLAELERKILKYVDYLENNSDVQSLKRS 266
QY 56 PEDILTKFI-----NCAKETHFAHMYKKNVNAE-NELKSGKIPGFRQMYTF 106
DB 267 KEDELKLNMLNELKSAEERDQL--ERKKNELKRTNELNELIK-- 312
QY 107 GDFRIFEGTDISSCRVYIKDTSQTKSL--GQATTEKGDTHIDNNK--KIQEMWTIH 161
DB 313 SDENDLQKQONSKSRKAKDELNELFTKFSNGSQSSAKEMELKMKNAIALEBELSTK 372
QY 162 GPKI--WEGMLCALTNGISESEKKNILQDYSYNNKLNNAEKDDCCLEKFAKPO---FLR 215
DB 373 NSOLIAKEGKLASIAQITQLESKLNRD--SOLGSEEE--LKTNRKLDKDIRAR 426
QY 216 WYVWSDEFCRE-RKKLEDDVEDVCIRAKDEGCGKNNKSN--NSCVAVCKEYEVYIRGKK 272
DB 427 EETYSKDEIRIIDLQKKYQLENDLFFVIKTHSEKTTITDNELESKDLIKILENDLVAQ 486
QY 273 TQYESQEGFENTERKRPKPEYN---SYSK-KDASEYLNDK 308
DB 487 EKYSKM-----EKELKEREFNKYSKLEDETTLINEK 520

RESULT 12
SMC2_YEAST
ID SMC2_YEAST STANDARD: PRT: 1170 AA.
AC P38989;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chromosome segregation protein SMC2 (DA-box protein SMC2).
GN SMC2 OR YFR031C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_Taxid=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95212908; PubMed=7698648;
RA Strunikov A.V., Hogan E., Koshland D.;
RT "SMC2, a Saccharomyces cerevisiae gene essential for chromosome
segregation and condensation, defines a subgroup within the SMC
family";
RL Genes Dev. 9:587-599(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;


```

RX MEDLINE-95400292; PubMed-7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasamura S.-I., Sasamura M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae.;
RN Nat. Genet. 10:261-268(1995).
RL [3]
RP SEQUENCE FROM N.A.
RC STRAIN-5288c / AB972;
RX MEDLINE-96287654; PubMed-8686381;
RA Eki T., Naitou M., Hagiwara H., Abe M., Ozawa M., Sasamura S.-I.,
RA Sasamura M., Tsuchiya Y., Shibata T., Watanabe K., Ono A.,
RA Yamazaki M.-A., Tashiro H., Hanaoka F., Murakami Y.;
RT Fifteen open reading frames in a 30.8 kb region of the right arm of
RT chromosome VI from Saccharomyces cerevisiae.;
RL Yeast 12:177-190(1996).
CC -1- FUNCTION: INVOLVED IN CHROMOSOME SEGREGATION IN MITOSIS. COULD BE
CC PART OF A CHROMOSOME CONDENSATION MOTOR.
CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH SMC1 OR OLIGOMERS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS
CC FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY.
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U05820; AAA17416.1; -
CC EMBL: D50617; BAA09270.1; -
CC PIR: S48530; S48530.
CC SGD: S0001927; SMC2.
CC InterPro: IPR003405; SMC_C.
CC InterPro: IPR003395; SMC_N.
CC DR Pfam: PF02463; SMC_N; 1.
CC DR Pfam: PF02483; SMC_C; 1.
CC DR ProDom: PD000006; ABC_Transport; 1.
CC KW Mitosis; ATP-binding; Coiled coil; Nuclear protein.
CC FT NP_BIND 32 39 ATP (POTENTIAL).
CC FT DOMAIN 172 469 COILED COIL (POTENTIAL).
CC FT DOMAIN 678 1027 COILED COIL (POTENTIAL).
CC FT DOMAIN 1093 1119 ALA/ASP-RICH (DA-BOX).
CC SEQUENCE 1170 AA; 133927 MW; 142841AAE109621F CRC64;
CC -----
Query Match 5.7%; Score 97.5; DB 1; Length 1170;
Best Local Similarity 22.1%; Pred. No. 14;
Matches 68; Conservative 50; Mismatches 118; Indels 71; Gaps 14;

```

```

QY 300 DASEYLK 306
Db 495 KNSEYLK 501
RESULT 13
CYAA-DICDI STANDARD; PRT; 1407 AA.
ID CYAA-DICDI
AC 003100;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Adenylate cyclase, aggregation specific (EC 4.6.1.1) (ATP
DE pyrophosphate-lyase) (Adenyl cyclase).
GN ACBA OR ACA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxId=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92233467; PubMed-1348970;
RA Pitt G.S., Milona N., Bortels J., Lin K.C., Reed R.R., Devreotes P.N.;
RT Structurally distinct and stage-specific adenyl cyclase genes play
RT different roles in Dictyostelium development.;
RL Cell 69:305-315(1992).
CC -1- FUNCTION: COORDINATES CELL AGGREGATION BY SYNTHESIZING THE CAMP
CC THAT INFLUENCES DIFFERENTIATION AND MORPHOGENESIS OF CELLS
CC WITHIN A DEVELOPING MULTICELLULAR STRUCTURE.
CC -1- CATALYTIC ACTIVITY: ATP -> 3',5'-cyclic AMP + diphosphate.
CC -1- ENZYME REGULATION: REGULATED BY A SURFACE RECEPTOR THROUGH A
CC GUANINE NUCLEOTIDE BINDING PROTEIN. BOTH POSITIVELY AND
CC NEGATIVELY REGULATED BY EXTRACELLULAR CAMP; THIS REGULATION IS
CC PART OF THE MECHANISM THAT ESTABLISHES THE OSCILLATORY CAMP WAVES
CC DURING AGGREGATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DEVELOPMENTAL STAGE: CELL AGGREGATION AND LATER STAGES.
CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO ADENYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L05499; AAA33163.1; -
CC EMBL: L05496; AAA33163.1; JOINED.
CC EMBL: L05497; AAA33163.1; JOINED.
CC EMBL: L05498; AAA33163.1; JOINED.
CC PIR: B42239; B42239.
CC DR HSP: P26769; IAB8.
CC DR Dictydb: IP001054; acaa.
CC InterPro: IPR001054; G_cyclase.
CC Pfam: PF00014; CYCC_2.
CC SMART: SM00044; CYCC_2.
CC PROSITE: PS00452; GUANYLATE_CYCLASES_1; 1.
CC PROSITE: PS50125; GUANYLATE_CYCLASES_2; 2.
CC Lysase; CAMP synthesis; Transmembrane.
CC KW Lysase; CAMP synthesis; Transmembrane.
CC FT DOMAIN 1 219 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 220 240 POTENTIAL.
CC FT TRANSMEM 244 264 POTENTIAL.
CC FT TRANSMEM 276 296 POTENTIAL.
CC FT TRANSMEM 304 324 POTENTIAL.
CC FT TRANSMEM 325 345 POTENTIAL.
CC FT TRANSMEM 353 373 POTENTIAL.
CC FT DOMAIN 374 962 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 963 979 POTENTIAL.
CC FT TRANSMEM 992 1012 POTENTIAL.
CC FT TRANSMEM 1018 1038 POTENTIAL.
CC FT TRANSMEM 1071 1091 POTENTIAL.

```

```

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ertlan K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fume S., Galizzi A., Gallenot N.,
RA Gilm S.V., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holstappel E., Hosono S., Hultio M.F., Itaya M., Jones L.,
RA Jorits B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,
RA Kobayashi Y., Koeter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapdus A., Lardinot S., Lauber J., Lazarevic C.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Pario V., Pohl T.M., Portecelle D., Porvillik S., Prescott A.M.,
RA Prescan E., Puig C., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekituchi J., Sekowska A., Seror S.J., Serrot R., Shin B.S., Soldo B.,
RA Sorochin M., Taccioni F., Takagi T., Takahashi H., Takematu K.,
RA Takenchi M., Tamakoshi A., Tanaka T., Terpsira P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Vlati A., Wamputt R., Wedler E., Wedler H., Weitzmeyer T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasmoco K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumslein E., Yoshikawa H., Zandhin A.,
RA "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [3]
RN CHARACTERIZATION.
RX MEDLINE-95318144; PubMed-7797557;
RA Perego M., Glaser P., Minutello A., Strauch M.A., Leopold K.,
RA Fischer W.;
RT "Incorporation of D-alanine into lipoteichoic acid and wall teichoic
RL acid in Bacillus subtilis: Identification of genes and regulation."
RL J. Biol. Chem. 270:15598-15606(1995).
CC -I- FUNCTION: COULD BE RESPONSIBLE FOR THE TRANSFER OF D-ALANINE
CC RESIDUES FROM D-ALA-UNDECAPRENOL PHOSPHATE TO THE
CC POLY(GLYCEROPHOSPHATE) CHAIN.
CC -I- PATHWAY: D-alanyl-lipoteichoic acid biosynthesis.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X73124; CAAS1558.1; -
DR EMBL: 299123; CAB15879.1; -
DR PIR: S39657; S39657.
DR Subtilisin; Bg10548; dltD.
KW Signal; Complete proteome.
FT SIGNAL 1 19
FT CHAIN 20 392
FT FT PROTEIN DLT.
SQ SEQUENCE 392 AA; 44608 MW; 0C0DDA4673028B86E CRC64;
-----
Query Match 5.7%; Score 97; DB 1; Length 392;
Best local similarity 22.2%; Pred. No. 4.4;
Matches 67; Conservative 36; Mismatches 115; Indels 84; Gaps 15;
-----
Db 18 PDMCKKNINDNSHSGACMPRRKCLCVDTGGELRKEP---DILTYFINCAKETHFA 74
128 PÖWIKRKGSDQH---FAPNSALQGLDLAFNDQI-DPEIKKKMMKRLRFAVQNDAI 182
-----
Db 75 WHKKKKKNVNAE---NELKSGKIPGCFKQWVYGGDRDIFFGDIDSS-CRYI---- 124
183 LSELKYKAMVNGQWKNALKRA-----AKYYSLEKDKDLYSTTSSEKKRITISOV 235
-----
Db 125 KDISQITKSLGDA---TTEKGDTHID-NKKLQEWMTIHGPKRIWGMICALTINGLSE 179
236 KDSWSLNLADQSGRRHSGSDNDEHIDNPVYKKLK-----PKV-----PKLKGK 282
-----
Db 180 SEKKNILODY-SYNNKLNNAEKDDCCIEFKASQDPIFPMWVMSQDCEAFKKV PKYVND 326
-----

```

```

Db      283 GRSAVSPSEYDEEMDILDKDAGAEPMFTIPVNGKMY-----321
QY      239 CIAKDEGCKNNKSNNSCVKCKEYENYITGKKTYEESQEGKFNTEKROKKEPYNYSK 298
Db      322 -----DTGPKKGRFTDYKRVNKQ-----IRAKGFOVADPSGH-----EYDYFM 362
QY      299 KD 300
Db      363 KD 364

```

```

RESULT 15
CYL2_HUMAN STANDARD; PRT; 348 AA.
ID CYL2_HUMAN
AC 014093:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
RT 15-JUN-2002 (Rel. 41, Last annotation update)
CYL2_C1 OR CYL2_C2 OR CYL2_C3
CYL2 OR CYL2

```

```

OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=95255491; PubMed=7737358;
RA Hess H., Heid H., Zimbelmann R., Franke W.W.;
RT "The protein complexity of the cytoskeleton of bovine and human sperm
RT heads: the identification and characterization of cyclin II."
RL Exp. Cell Res. 218:174-182(1995).
CC -!- FUNCTION: POSSIBLE ARCHITECTURAL ROLE DURING SPERMATOGENESIS. MAY
CC BE INVOLVED IN SPERMATID DIFFERENTIATION.
CC -!- SUBCELLULAR LOCATION: CALYX; SPERM HEAD CYTOSKELETAL STRUCTURE.
CC -!- TISSUE SPECIFICITY: TESTIS.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z46788; CA86752.1; -
GeneW; HGNC:2583; CYL2.
MIM; 604035; -

```

```

CYTOSKELETON: Structural protein; Repeat; Sperm; Spermatogenesis.
FT DOMAIN 25 347 31 X 3 AA REPEATS OF K-K-X.
FT REPEAT 157 240 3 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 157 184 1.
FT REPEAT 185 212 2.
FT REPEAT 213 240 3.
SQ SEQUENCE 348 AA; 39079 MW; D86766599C1809E7 CRC64;

```

```

Query Match 5.7%; Score 96.5; DB 1; Length 348;
Best Local Similarity 23.8%; Pred. No. 4.2;
Matches 72; Conservative 33; Mismatches 104; Indels 93; Gaps 13;

```

```

QY 38 RROKL-----CVRDLTGGEL-RRPEDILTFFINCAKETHFAMHKYKKDNVNAENELKSG 92
Db 95 RROPLKPTFVEVDKSAKEIGKGEDKT-----OKDTTDSSEELKOG 137
QY 93 KIPGFRKQMYVFFGDFRIDFFGTDISSCRYIK-DTSOTIK-----SKLDQAT---TE 142
Db 138 K-----KDSKKGKIDIEKKEKLDARKKSKGKDAEKGKDSATSESEDE 181
QY 143 KGDTHIDNKKLQEMWTFIHGPKIWESGLCALTNGL-----SESKNIILODYSYNKLNNA 197
Db 182 KGAKKDNKKDKKD-----SNKGKDSATSESEGEKGTGERKSKKGRKDSK 225

```

```

QY 198 EKDDCCLEKFAKPOFLRMVYVSEDFCRBRKRLKEDKVEDVCIKAKDYEGCKNNKSNNSC 257
Db 226 KGDSDAIELOAVKAD-----EKKDEGCKKDKANKGDSKAKKDAKFKTKGKKDKKPPSS 279
QY 258 V-----KCKEYE-----NYTGKKTYEESQEGKFNTEKROKKEPYNYSKQDASEY 304
Db 280 TDSKSDVAKKESKDKATKDAKKVAKKDTESADSKKDAKKNAKND-----AKKDAKKN 334
QY 305 LK 306
Db 335 AK 336

```

Search completed: June 20, 2003, 15:03:34
Job time : 16.5539 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 14:54:09 ; Search time 48.1396 Seconds
(Without alignments)
1318.301 Million cell updates/sec

Title: US-10-087-013-11

Perfect score: 1706

Sequence: 1 GNDGSENEISGNCNPKESYPDW.....KKREYNSYSKKDASEYLKDK 308

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

al number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_21:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhcc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.potent:*
13: sp.virus:*
14: sp.unclassified:*
15: sp.rvtrus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1706	100.0	2710	5	Q9XZB8 plasmodium
2	894	52.4	3006	5	Q26032 plasmodium
3	669.5	39.2	3078	5	Q26031 plasmodium
4	617.5	36.2	3287	5	Q8T326 plasmodium
5	600	35.2	3026	5	Q26030 plasmodium
6	572	33.5	3542	5	Q9U5M2 plasmodium
7	571.5	33.5	2658	5	Q8T5G0 plasmodium
8	569.5	33.4	1615	5	Q8T325 plasmodium
9	526	30.8	2706	5	Q15870 plasmodium
10	521	30.5	921	5	Q25989 plasmodium
11	515.5	30.2	2527	5	Q95W83 plasmodium
12	444	26.0	1685	5	Q9U4A2 plasmodium
13	437	25.6	2209	5	Q9U0G6 plasmodium
14	433	25.4	1711	5	Q96108 plasmodium
15	407.5	23.9	2169	5	Q97312 plasmodium
16	395	23.2	2042	5	Q25766 plasmodium

17	386	22.6	2228	5	Q60991 plasmodium
18	385.5	22.6	2277	5	Q9U0G5 plasmodium
19	383	22.5	2212	5	Q94657 plasmodium
20	375	22.0	184	5	Q9BUE6 plasmodium
21	371.5	21.8	2209	5	Q97324 plasmodium
22	368.5	21.6	195	5	Q00831 plasmodium
23	367.5	21.5	2182	5	Q26034 plasmodium
24	362	21.2	176	5	Q9BJR0 plasmodium
25	361.5	21.2	2647	5	P90580 plasmodium
26	354	20.8	2135	5	Q61077 plasmodium
27	354	20.8	2163	5	Q9NEB6 plasmodium
28	352	20.6	2664	5	Q26033 plasmodium
29	342.5	20.1	2197	5	Q96296 plasmodium
30	333	19.5	182	5	Q00832 plasmodium
31	327.5	19.2	616	5	Q8T9M3 plasmodium
32	327.5	19.2	616	5	Q8T9L8 plasmodium
33	327.5	19.2	1210	5	Q96774 plasmodium
34	327.5	19.2	1210	5	Q8WS31 plasmodium
35	325.5	19.1	177	5	Q9BUE8 plasmodium
36	321	18.8	162	5	Q9BUE7 plasmodium
37	311	18.2	2924	5	Q25733 plasmodium
38	298.5	17.5	1729	5	Q25734 plasmodium
39	286	16.8	173	5	Q25988 plasmodium
40	268.5	15.7	699	5	Q25987 plasmodium
41	265.5	15.6	2647	5	Q9U4X0 plasmodium
42	265	15.5	420	5	Q9BHE9 plasmodium
43	265	15.5	420	5	Q9BHE2 plasmodium
44	263	15.4	169	5	Q25990 plasmodium
45	263	15.4	420	5	Q9BMY4 plasmodium

ALIGNMENTS

RESULT 1
Q9XZB8 PRELIMINARY; PRT; 2710 AA.
AC Q9XZB8:
DT 01-NOV-1999 (TREMBLREL. 12, Created)
DT 01-NOV-1999 (TREMBLREL. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLREL. 19, Last annotation update)
DE Variant-specific surface protein.
GN VAR.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
CX NCBI_TaxID=5633;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS2;
RX MEDLINE=99238507; Pubmed=10220443;
RA Reeder J.C., Cowman A.F., Davern K.M., Beeson J.G., Thompson J.K.,
Rogerson S.J., Brown G.V.;
RT "The adhesion of Plasmodium falciparum-infected erythrocytes to
RT chondroitin sulfate A is mediated by P. falciparum erythrocyte
RL Proc. Natl. Acad. Sci. U.S.A. 96:5198-5202(1999).
DR EMBL: AF134154; AAD29126.1;
DR InterPro: IPR004258; PfEMP.
DR Pfam: PF03011; PfEMP; 2.
SQ SEQUENCE 2710 AA; 309247 MW; 0AC64F0A3D5BF512 CRC64;

Query Match 100.0%; Score 1706; DB 5; Length 2710;
Best Local Similarity 100.0%; Pred. No. 1,1e-115;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNDGSENEISGNCNPKESYPDWCKRNIDNSHGACMPRRKLCVRLDLOGEIRKPEDIL 60
DB 908 GNDGSENEISGNCNPKESYPDWCKRNIDNSHGACMPRRKLCVRLDLOGEIRKPEDIL 967
QY 61 TKFLNCAKETHFMHKKYKKNVNAENELKSGKIPGEFRKMYTTTGFDFRIFFGTIDSS 120
DB 968 TKFLNCAKETHFMHKKYKKNVNAENELKSGKIPGEFRKMYTTTGFDFRIFFGTIDSS 1027

```

QY 121 CRYIKDTSQIKSKLDQATEKGDTHIDNKKLOEMWTIHGKIEWGMICALTNGLSSES 180
DB 1028 CRYIKDTSQIKSKLDQATEKGDTHIDNKKLOEMWTIHGKIEWGMICALTNGLSSES 1087
QY 181 EKKNILODYSYNKLNNAEKDDCCLEFASKPOFLRMVEMSDPEFCERKLEKVDVCI 240
DB 1088 EKKNILODYSYNKLNNAEKDDCCLEFASKPOFLRMVEMSDPEFCERKLEKVDVCI 1147
QY 241 KADYEGCKNNKSNNSCVKCYEYENYITGKTOYESQEGKFTEROKKPEYNSYKRD 300
DB 1148 KADYEGCKNNKSNNSCVKCYEYENYITGKTOYESQEGKFTEROKKPEYNSYKRD 1207
QY 301 ASBYLKDK 308
DB 1208 ASBYLKDK 1215

```

RESULT 2

```

Q26032 PRELIMINARY; PRT; 3006 AA.
Q26032; 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE Variant-specific surface protein.
RT cytoadherence and antigenic variation of Plasmodium falciparum-
RT infected erythrocytes."
RN Cell 82:89-100(1995).
RA EMBL; L40609; AAA75397.1; -
DR InterPro; IPR004258; PfEMP.
DR Pfam; PF03011; PfEMP; 2.
SQ SEQUENCE 3006 AA; 343769 MW; 8FD5FF475F943C74 CRC64;

```

```

Query Match 52.4%; Score 894; DB 5; Length 3006;
Best Local Similarity 53.0%; Pred. No. 1.7e-56;
Matches 172; Conservative 38; Mismatches 91; Indels 12; Gaps 4;

```

```

5 SNEISGCM-----KESYPMDCCKNNIDNSHGACMPPRROKLCVRODLOGEIRKPEDI 59
1236 TDDIDGCKQKAKGDKRPGMDCNSQIHTHNGACMPPRROKLCVSGLTDRKKAIEYI 1295
QY 60 LTKFLINCAKETHFAMHKYKKNVNAENELSGKIPGFRKQMYTFEGDFRDIFFGTDIS 119
DB 1296 RTEFIKSAIETHFAMDRYKEDNGEAEALKNGLNIPBEGRKQMYTFEDDYDIFGRDIS 1355
QY 120 SCRYIKDTSQIKSKLDQATEKGDTHIDNKKLOEMWTIHGKIEWGMICALTNG 176
DB 1356 THAVISGVSPKVTITLEKENAKYAAKONS---NNEILDMDWDHGHKDIWEGMICALTHK 1412
QY 177 LSESE-KKNILODYSYNKLNNAEKDDCCLEFASKPOFLRMVEMSDPEFCERKLEKDV 235
DB 1413 ISDEKKEKKEIKNNKSYKLNKSPGSKVEDEFAKPOFLRMVEMSDPEFCERKLEKDV 1472
QY 236 EDVCIAKADYEGCKNNKSNNSCVKCYEYENYITGKTOYESQEGKFTEROKKPEYNS 295
DB 1473 KVSQSDAKDYDGCNKNTKSNASCVCAYVEDYITKKVEYETKQKGRDAEKITTEGEG 1532
QY 296 YSKKDASEYLKDK 308
DB 1533 FSTKDASEYLKDK 1545

```

RESULT 3

```

Q26031 PRELIMINARY; PRT; 3078 AA.
Q26031; 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE Variant-specific surface protein.
RN Cell 82:89-100(1995).
RA EMBL; L40608; AAA75396.1; -
DR InterPro; IPR004258; PfEMP.
DR Pfam; PF03011; PfEMP; 2.
SQ SEQUENCE 3078 AA; 349297 MW; C8037C2BC3CD7C3 CRC64;

```

```

Query Match 39.2%; Score 669.5; DB 5; Length 3078;
Best Local Similarity 40.9%; Pred. No. 3.8e-40;
Matches 133; Conservative 54; Mismatches 109; Indels 29; Gaps 9;

```

```

1 GNDGSNEISGCM-----KESYPMDCCKNNIDNSHGACMPPRROKLCVRODLOGEIRKPE 57
1356 GKNQITTYGCEGCKPESYPMDCCKNNIDNSHGACMPPRROKLCVLYIAHESOTENIKTDD 1415
QY 58 DILTKFLINCAKETHFAMHKYKKNVNAENELSGKIPGFRKQMYTFEGDFRDIFFGTDIS 117
DB 1416 NIKDAFIKTAAEFTLSWQYKSKNDSEAKITLDGLIPSGFLRSMWTFGDRICLNTD 1475
QY 118 ISSCRYIKDTSQIKSKLDQATEKGDTHIDNKKLOEMWTIHGKIEWGMICAL 172
DB 1476 ISKNO--NDVAKA-KDKIGKFFSKDGSKSPGSLR-----QEMWKTNGPEIKWGMICAL 1525
QY 173 LTNGLSESE-KKNILODYSYNKLNNAEKDDCCLEFASKPOFLRMVEMSDPEFCERK 231
DB 1526 LTKVYTDIDNKKRINKNDYSYKVNQSONGNLFPAKPOFLRMVEMSDPEFCERK 1585
QY 232 EDKVEDVCIAKADYEGCKNNKSNNSCVKCYEYENYITGKTOYESQEGKFTEROK 289
DB 1586 ENIKIDACNEINSTQOC--NDAKHRCQACRAYOYEYENKKEFSGQTNNFVLKANVOPQ 1643
QY 290 KPEYNSYKRD-----ASEYLKDK 308
DB 1644 DPEYKGYEYKDGVOPIQONEYLLOK 1668

```

RESULT 4

```

Q8T326 PRELIMINARY; PRT; 3287 AA.
Q8T326; 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE Variant-specific surface protein.
RN TMT80VAR2.
RA EMBL; L40609; AAA75397.1; -
DR InterPro; IPR004258; PfEMP.
DR Pfam; PF03011; PfEMP; 2.
SQ SEQUENCE FROM N.A.
RN STRAIN-TM180;
RX MEDLINE-21927235; PubMed-11930336;
RA Rowe J.A., Kyes S.A., Rogerson S.J., Babiker H.A., Raza A.;

```

RT "Identification of a conserved Plasmodium falciparum var gene implicated in malaria in pregnancy."
 RT J. Infect. Dis. 185:1207-1211(2002).
 DR EMBL: AJ420411; CAD20867.1; -
 FT NON_TER 3287 3287
 SQ SEQUENCE 3287 AA; 588F866FC244536 CRC64;
 Query Match 36.2%; Score 617.5; DB 5; Length 3287;
 Best Local Similarity 41.7%; Pred. No. 2,5e-36;
 Matches 130; Conservative 56; Mismatches 101; Indels 25; Gaps 11;
 OY 8 ISGCMFKRS--YPMDCCKNIDNS-HSGACMPRRQKLCVRL--TGGGEIRKP-DILT 61
 Db IEGCMFKRNGYKPKMEC--NVCKSKENACMPRRKIKCINNIQYLNKTEKPEKIDKN 2138
 OY 62 KFINCAKEHPAMHKYKKNVNAENELSKIPGCFKQMYTTCGDRDIFGDISC 121
 Db AFICGALEIFYFWLYKLDNPTAENELKNGRIPEEFKIMYYTGDKDMFEGDITNE 2198
 122 RYIKDTSOTIKSLGDAQTEKGDPIHIDNKKIQEWMTHGPKIMEGMLCALTNGLS-ES 180
 2199 SKIITINNNTYIILKNENKKKQDKRTDELLKNFWEENKKLIMEGMILGTLHTLDET 2258
 OY 181 EKKNILODYTNKLNNAEKDDCCLEKFAKPOFLRWYVWSDFCRERKKLEKVEDYCI 240
 Db EKKKTTDSEYTDIKLSPS--LEEFAPKRPQFLRWFTWMAEEFCNKRK--OLENLKE 2312
 OY 241 KAKDEGCKNNKNSCVKCEYENYITGKKTOYESOGCKNTEKROKKPEYNSY---- 296
 Db 2313 KCPDVT-CSVDSTKKECDSCCKKYOTFIKQMTQYERKGRKNDKDTDK--YKDIPIES 2369
 OY 297 ---SKKDASEYL 305
 Db 2370 YIDNATDAHEYL 2381
 RESULT 5
 OQ 026030 PRELIMINARY; PRT; 3026 AA.
 AC 026030;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1998 (TREMblrel. 08, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE Variant surface protein (Fragment).
 GN VAR.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_Taxid=5833;
 [11]
 SEQUENCE OF 2044-2922 FROM N.A.
 RC STRAIN-IT 4/25/5;
 RA MEDLINE-95330813; PubMed-7606788;
 RA Su X.Z., Heatwole V.M., Wertheimer S.P., Guinet F., Herfelde J.A., Peterson D.S., Ravetch J.A., Wellens T.E.;
 RA "The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of Plasmodium falciparum-infected erythrocytes."
 RT Cell 82:89-100(1995).
 RL [12]
 RN SEQUENCE OF 2044-2922 FROM N.A.
 RC STRAIN-IT 4/25/5;
 RA Smith J.D., Chlitis C.E., Craig A.G., Roberts D.J., Hudson-Taylor D.E., Peterson D.S., Planches R., Newbold C.I., Miller L.H.;
 RA "Switches in the expression of Plasmodium falciparum var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes."
 RT Cell 0:0-0(1995).
 RL [13]
 RN SEQUENCE FROM N.A.
 RC STRAIN-IT 4/25/5;
 RX MEDLINE-99094502;
 RX Smith J.D., Kyes S., Craig A.G., Fagan T., Hudson-Taylor D.,

RA Miller L.H., Baruch D.I., Newbold C.I.;
 RT "Analysis of adhesive domains from the AVAR Plasmodium falciparum erythrocyte membrane protein-1 identifies a CD36 binding domain."
 RT Mol. Biochem. Parasitol. 97:133-148(1998).
 RN [14]
 RN SEQUENCE FROM N.A.
 RC STRAIN-IT 4/25/5;
 RA Smith J.;
 RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
 RN [15]
 RN SEQUENCE FROM N.A.
 RC STRAIN-IT 4/25/5;
 RA Kyes S., Smith J.;
 DR Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: I42244; AA03351.1;
 DR InterPro: IPR004258; PfEMP.
 DR Pfam: PF03011; PfEMP. 2.
 FT NON_TER 3026 3026
 SQ SEQUENCE 3026 AA; 348124 MW; EC0A971A82193788 CRC64;
 Query Match 35.2%; Score 600; DB 5; Length 3026;
 Best Local Similarity 35.9%; Pred. No. 4.3e-35;
 Matches 122; Conservative 58; Mismatches 112; Indels 48; Gaps 9;
 OY 1 GNDGSNEISGCN----PREYPMDCCKNIDNSHSGACMPRRQKLCVRLDTGGEIRK 55
 Db 1996 GDDGTRKRIIECCTKTYPTKNDYFGNNCTQKVINREGSCMPRRQKLCIHLNLESEKAT 2055
 OY 56 PEDILTKFINCAKETHPAMHKYKKNV-----NAENELSKRIPGFK 100
 Db 2056 ETELKRAFECAAIETFWLMDKYEKDEKTEGGGISDDDDPOKKLEGSTIPEDFKR 2115
 OY 101 QMYTFGDFRDIFFGNDISSCXYINDTSOTIKSLGDAQT--EKGDPIHIDNKKIQEWM 158
 Db 2116 QMYTFGDFRDLFGTIDSK-----CHGESALGKIDSLFKNGDQSPSGKTPTEWM 2168
 OY 159 TTHGPKIMEGMLCALT---NGLSESEKKNILQDYSYNKLNAEKDDCCLEKFAKPOFL 214
 Db 2169 NNYGPDIMKGVCGLSHHIKNCKNEDQLRKNLIDNNKYTKISS-----LEDFASAPQFL 2222
 OY 215 RWYVWSDFCRERKKLEKVEDVCIKADYE-GCKNNKSNNSCVKCEYENYITGKKT 273
 Db 2223 RWFIEWGDFCRERVVKINQLTGC--NEYECGSGENCKKACKNACRAKSWLKDWD 2279
 OY 274 QYESOGCKNTEKROK-----PEYNSYKKNASSTYLNK 308
 Db 2280 QYEQQTAFFDKDKDKKEDGTSAEVDVAVSSVHEYLDE 2319
 RESULT 6
 OQ 0905M2 PRELIMINARY; PRT; 3542 AA.
 AC 0905M2;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
 DE FC3 CSA ligand (Fragment).
 GN VAR.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 NCBI_Taxid=5833;
 [11]
 RN SEQUENCE FROM N.A.
 RC STRAIN-FC3;
 RX MEDLINE-20006305; PubMed-10535993;
 RA Butfer P., Gamain B., Scheidig C., Baruch B., Oishi S., Fujii N., Fusaai T., Parzy D., Miller L.H., Gysin J., Scherf A.;
 RT "Plasmodium falciparum domain mediating adhesion to chondroitin sulfate A: A receptor for human placental infection."
 RT Proc. Natl. Acad. Sci. U.S.A. 96:12743-12748(1999).
 RL EMBL: AJ13811; CAB59840.1; -
 DR InterPro: IPR001219; Neurotoxin.
 DR InterPro: IPR004258; PfEMP.

DR Pfam; PF03011; PFEMP; 1.
DR PRINTS; PR00284; TOXIN.
FT NON_TER 3542 3542
SQ SEQUENCE 3542 AA; 413089 MW; 970DB5EE8DA2EC2 CRC64;

Query Match 33.5%; Score 572; DB 5; Length 3542;
Best Local Similarity 39.9%; Pred. No. 5.6e-33;
Matches 131; Conservative 54; Mismatches 105; Indels 38; Gaps 14;

QY 3 DGSNEISGCPKPE--SYPPDMD--KKININSHGACMPPRROKLCVARD--LTQGEIRK 55

DB 2092 DSRGGLEGCPKPYGQYPRMGCTGVGSKEN--ENGICMPPRKRLCTINNLYLVETENKR 2150

QY 56 PEDILKFTINCAKETHFAHMKYKKDNVAENELKSGKIPREGKQMYTFGDFRDIFFG 115

DB 2151 DNDIKAFIKCAIEIETQFLMKLYIENPAENLQNGTIDPEKRLIMYYGYGKMFEG 2210

QY 116 TDISSCRITKDTSTQTKSKIGDQATEKGDTHIDNKKLOE---WMTIHPKIMEGMLC 171

DB 2211 TDISNDKKTITVNSVTTLINENKKRQ-----DKKDELRKIFWEKKKFIWEGMITY 2264

QY 172 ALTNGLS-ESEKKNILQDYSYNKLNNNAEKDDCLEKFAKPOFLRRYVENSDEFCEKRRK 230

DB 2265 GLTYHLTDENEKKEKIDNTQYNDMTKLTPS--LEEFVVRKPOFLRMTTEVAEEFCCKRKE 2321

QY 231 LEDKVEDVCIKADYEGC--KNNKSNNSGVKVEKENYITGKKTQYESQEGKFTNEKRQ 288

DB 2322 QLLKLKAGC---KEYE-CNGSNDGKTQEOCAEACVTONFKTKKTEYERREKFKDKG 2377

QY 289 KRPEVNSY--SKKD-----ASEYLKDK 308

DB 2378 KR--YKDYPTSTERDIEKATCAHEYLNMK 2403

RESULT 7

ID 08T5G0 PRELIMINARY; PRT; 2658 AA.

AC 08T5G0; 01-JUN-2002 (Tremblrel. 21, Created)

DT 01-JUN-2002 (Tremblrel. 21, last sequence update)

DE Erythrocyte membrane protein-1 (Fragment).

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;

RN [1]

PD SEQUENCE FROM N.A.

STRAIN=JDP8;

RT Chatopadhyay R., Pillai C.R., Chitnis C.;

RL "Identification of a domain responsible for binding to intercellular

adhesion molecule-1 from a Plasmodium falciparum field isolate."

DR EMBL; AT028643; AAK49742.1; -

FT NON_TER 2658 2658

SQ SEQUENCE 2658 AA; 301326 MW; AEE33BC9167E7B7 CRC64;

Query Match 33.5%; Score 571.5; DB 5; Length 2658;

Best Local Similarity 38.0%; Pred. No. 4.4e-33;

Matches 127; Conservative 54; Mismatches 102; Indels 51; Gaps 12;

QY 2 NDGSNEISGCPKPE--SYPPDMD--KKININSHGACMPPRROKLCVARDLTQGEIRK 55

DB 1317 DDFKQVGGCNAKIKUNESYPMWTCVNSKFENNEGPCMPRRQKLYLKELEND 1376

QY 56 PEDILKFTINCAKETHFAHMKYKKDNVAENELKSGKIPREGKQMYTFGDFRDIFFG 115

DB 1377 EOKFKDAFIKTAAEFTLSQYKSKNSMDIKLQSGEIPDEELRSMYFYGDYDLCN 1436

QY 116 TDISSCRITKDTSTQTKSKIGDQATEKGDTHIDNKKLOEWMTHHPKIMEGMLC 169

DB 1437 TDIS-----KKEGD-VSDAKGKIDAFKNTYDITNR--TKWMDTNGPEIWEGA 1480

QY 170 LCATNLGSESE--KKNILQDYSYNKLNNNAEKDDCLEKFAKPOFLRRYVENSDEFCEKRR 228

DB 1481 LCATLHGVIINTDMKRRKIKTDYSYKL--QSKNVTPLKFAERQFLRMILWGGDDCRQ 1539

QY 229 KRLKEDVEDVCIKADYEGCCKNNKSNNS-----CYKCKEYENYITGKKTQYESQ 279

DB 1540 KKKYNELKCKCNK-----CNGNNTSDECKTKVECCOKKEGYKFTTEQEMNNKOR 1593

QY 280 GKFTNEKRQ--KRPEVNSYKDAE-----LYLKD 307

DB 1594 NKETLYTOVKTSRSSTISSDPTETKLKLYNE 1627

RESULT 8

ID 08T325 PRELIMINARY; PRT; 1615 AA.

AC 08T325; 01-JUN-2002 (Tremblrel. 21, Created)

DT 01-JUN-2002 (Tremblrel. 21, last sequence update)

DE Erythrocyte membrane protein-1 (Fragment).

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;

RN [1]

PD SEQUENCE FROM N.A.

STRAIN=TM284;

RT MEDLINE=21927235; PubMed=11930336;

RA Rowe J.A., Kyes S.A., Rogerson S.J., Babiker H.A., Raza A.;

RT "Identification of a conserved Plasmodium falciparum var gene

implicated in malaria in pregnancy."

J. Infect. Dis. 185:1207-1211(2002).

DR EMBL; AJ420412; CAD20868.1; -

FT NON_TER 1615 1615

SQ SEQUENCE 1615 AA; 188639 MW; FCA896C00B86DEA6 CRC64;

Query Match 33.4%; Score 569.5; DB 5; Length 1615;

Best Local Similarity 37.3%; Pred. No. 3.6e-33;

Matches 125; Conservative 57; Mismatches 104; Indels 49; Gaps 14;

QY 1 GNDGSNEISGCPKPE--SYPPDMDCKKNINSHGACMPPRROKLCVARDLTQGEIRK 58

DB 1253 GNGGTTKVGECNVKTGGEYPEWCEKKNINSHGACMPPRRKLCTVHFLKE--LKVETDCK 1311

QY 59 ILTKFTINCAKETHFAHMKYKKDNVAE--NELKSGKIPREGKQMYTFGDFRDIFFG 115

DB 1312 LREAFIQCSAETFLMKTKREDNNGEDLQNLQESGILIPDFKROMETTFGDRDLG 1371

QY 116 TDISSCRITKDTSTQTKSKIGDQATEKGDTHIDNKKLOEWMTHHPKIMEGMLC 172

DB 1372 KDIGS--DVGNNQKINSALKTIGQ-----SDEKRRNMWNGIKEDWKGKMGVC 1418

QY 173 LINGLSESEK--KNILQD-----YSYNKLN--NAEKDDCLEKFAKPOFLRRYVENS 222

DB 1419 LSHDVGNNHKEIYKRRKIMEDPONNKYQYNSVKTDEPSCGTLSEFAKVQOFLMLTEWMD 1478

QY 223 EFCREKRLKEDVEDVCIKADYEGCCKNNKSNNSGVKVEKENYITGKKTQYESQEGK 282

DB 1479 DYCHTQKLYKEVESTC--KSNQDLK-----DTECKKKCEDEYKMKKKKEWILQD--KY 1531

QY 283 NTEKROKPEYN-----SYSKKDASEYLKDK 308

DB 1532 YKDERDK--RPNQHGAVKDYTGINAIDYLRK 1565

RESULT 9

ID 015870 PRELIMINARY; PRT; 2706 AA.

AC 015870; 01-JAN-1998 (Tremblrel. 05, Created)

DT 01-JAN-1998 (Tremblrel. 05, last sequence update)

DE 01-DEC-2001 (Tremblrel. 19, last annotation update)


```

QY 166 MEGMICALTNGLSESEKKNIIOD---YGVNKLNNNAEKDDCCLEKFAKSPQFLRMVEMSD 222
DB 1427 MEGMICALTNGLSESEKKNIIOD---YGVNKLNNNAEKDDCCLEKFAKSPQFLRMVEMSD 1483
QY 223 EFCRRERKLEKVEDVC---IRAKDYEG-C-KNNKSNNSCVKCEYEVNIGKKTQYSSQ 278
DB 1484 HFCRRERKLEKVEDVC---IRAKDYEG-C-KNNKSNNSCVKCEYEVNIGKKTQYSSQ 1543
QY 279 EGFTEKERRKKRPPVNSYSKSDASYLKD 308
DB 1544 KVFKEKTEKGYNDPDTIQTSTAYEYELKK 1573

RESULT 12
QY 0904A2 PRELIMINARY: PRT: 1685 AA.
AC 0904A2:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
VAR. Variant surface protein PfEMP1 (Fragment).
GN Plasmodium falciparum.
OS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN-IT:
RX MEDLINE=20144115; PubMed=10677532;
RA Smith J.D., Grais A.G., Kiehl N., Hudson-Taylor D., Kyes S., Fagen T.,
RA Pitches R., Baruch D.I., Newbold C.I., Miller L.H.,
RT Identification of a Plasmodium falciparum intercellular adhesion
RT molecule-1 binding domain: A parasite adhesion trait implicated in
RT cerebral malaria.
RT Proc. Natl. Acad. Sci. U.S.A. 97:1766-1771(2000).
DR EMBL: AF193424; AAF1980.1;
DR InterPro: IPR000454; Eub. ATPase_Csub.
DR InterPro: IPR004258; PFEMP.
DR Pfam: PF03011; PFEMP; 1.
DR PROSITE: PS00605; ATPase_C; 1.
FT NON TER 1685 1685
SQ SEQUENCE 1685 AA; 193671 MW; D1PD426666B0551E CRC64;

Query Match 26.0%; Score 444; DB 5; Length 1685;
Best Local Similarity 34.4%; Pred. No. 5.2e-24;
Matches 115; Conservative 46; Mismatches 111; Indels 42; Gaps 12;

DB 6 NEISGCMKESYPMDCCKKINDSHSGACMPRRQKLCVARDLTOGGEIRKPEDILTKFIN 65
DB 1239 NNIHNCCKTBD-AKWKCENTRLGDEGVCMPPRRONLCVHYLTFLINDSKKEEDLREAFIK 1297
QY 66 CAKETHTFAWHKKYKDVNNAENELKSGKTEGFRKQYVTFGDPDRDIFFTDSSCYIK 125
DB 1298 SAAAFETLLQOYNSKRVEDDKILHRMIPPEFRSMFTFGDRLDCLDPTDISEKTADH 1357
QY 126 DTSSOTISKLGDDATTEKGDTHIDNKKLOE---WMTIHGPKIWEGMICALTNGLSESEK 182
DB 1358 DVT-TAKKTI--TAVFOKIGSKTTNGKLVLEPREGWMEYGLSTIMKGLCALSTN-TEKK 1413
QY 183 -----KNIIDQSYNKLNNNAEKDDCCLEKFAKSPQFLRMVEMSD---LED 233
DB 1414 MDEGVRYLIMKYIK--NNDIKI--YLEEFASRPPLRMVTEWGEDFVKNRKKEIVSLK 1469
QY 234 KVEDVCIK--AKDYEGCKNNKSNNSCVKCEYEVNIGKKTQYSSQEGKFE----- 282
DB 1470 KCDSCHLRNNGTSNKTCDNENCGACTQCEKTKKMMERKKHYSSOKKFFOLYKNSATY 1529
QY 283 -----NTEKRRKKPEY-NSYSKSDASEYIK 306
DB 1530 NNGIAYEANSFTYKNDPEVTEANSKAKHARDYIK 1563

```

```

RESULT 13
QY 0900G6 PRELIMINARY: PRT: 2209 AA.
AC 0900G6:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Val, MALP2.56.
GN MALP2.56, VAR.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN-3D7;
RA Devlin K., Bowman S., Churcher C., Harris B., Harris D., Lawson D.,
RA Quail M., Barrell B.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL035475; CAB62897.1;
DR InterPro: IPR004258; PFEMP.
DR Pfam: PF03011; PFEMP; 2.
SQ SEQUENCE 2209 AA; 250778 MW; 7A349F6FEFE9F7CB CRC64;

Query Match 25.6%; Score 437; DB 5; Length 2209;
Best Local Similarity 28.2%; Pred. No. 2.2e-23;
Matches 111; Conservative 51; Mismatches 114; Indels 118; Gaps 14;

QY 1 GNDGSEISGCMKPK-----ESYPMDC-----KKNIDSHSGACMPRRQKLCVARDL 47
DB 923 GNFGD---ACGTIKYDKYGRKPKNCKIPSGDKTAPSSDSNGSCVPPRRRLVYGGI 979
QY 48 TO-----GGERKPEDILTK-----FINCAKETHTFAWHKKYK-----81
DB 980 TKWASGNTVVSQAQPPQGGTSPDNKLRDAFIOSALETFFLMKMKDKIEKKOOO 1039
QY 82 -----NVNNEELK---SGKIEGFRKQYVTFGDPDRDIFFTDSSCYIKINDISO 129
DB 1040 KNLVANTSVNGKHEQKLEQSSILPEDRKQMFYLGDRDLDFKDISGDKNMDPTIE 1099
QY 130 TTSKLGDDATTEKGDTHIDNKKLOEWMTHGPKIWEGMICALT-----174
DB 1100 KINGILPKNGTSPSPAKNKP-----TEWWSONGEHIMNMICALTYDVTWASGDKPTQNE 1154
QY 175 -----NGLSESEKKNIIODSYN-----KLNNAEKDDCC-----LEKFAKSPQFL 214
DB 1155 KVEKALWDEONNKPKNVDYQSSVTTGEGAEGLQSTDSKDAARGKPTLIDSTIKRPYF 1214
QY 215 RMVEMSDSEFCRRERKLEKVEDVCIRAKDYEG-----CKNNKSN-----255
DB 1215 RYLEWGNQNFCKERRK---RLNDIKYECRGDENITRYGSGYGEDCKNNLPENPSTFKDLE 1271
QY 256 --SCVAVCKEYEVNIGKKTQYSSQEGKFNTEKR 287
DB 1272 YPTCAVYCRFYKMKMINTKTEYKQEKIYVQOKK 1305

RESULT 14
QY 096108 PRELIMINARY: PRT: 1711 AA.
AC 096108:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DGC-2001 (TREMBLrel. 19, Last annotation update)
DE PFEMP1.
GN PF0010W.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN 111
RP SEQUENCE FROM N.A.
RC MEDLINE=99021743; PubMed=9804551;
RX Gardner M.J., Tetteelin H., Carucci D.J., Cummings L.M., Aravind L.,
RX Koonin E.V., Shallow S., Mason T., Yu K., Fujii C., Pederson J.,

```

RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
 RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
 RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
 RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
 falciparum";
 RL Science 282:1126-1132(1998).
 DR EMBL: AE001366; AAC71792.1; -
 DR InterPro: IPR004258; PFEMP.
 DR Pfam: PF03011; PFEMP; 1.
 SO SEQUENCE 1711 AA; 196550 MW; 7BBD6C52742235B CRC64;

Query Match 25.4%; Score 433; DB 5; Length 1711;
 Best Local Similarity 33.4%; Pred. No. 3.3e-23;
 Matches 104; Conservative 49; Mismatches 110; Indels 48; Gaps 15;

DB 20 WDC--KKNDNSHSGA-CMPRRKQKIC-----VRPLTGGGELTRKPEDITKINCAAKE 70
 856 WRCIAPSGTTSGDGAICVPRTOELCYLKLKLSPTQKG-----LEAETKTRADE 908
 71 TRFAHKKYKKNV-----AENELSGKRIEGRKQMYTTEGDFRIFGTDIS 119
 909 TYLLMKYKEDKQNETASTELDIDDPOTLNGEITPDRKQMYTFYFGDRLFLG---- 964
 120 SCRTYKDTSGTTSKLGDOATTEKGDTHIDNKKL---OEWMTIHGPKIWESMLCALIN 175
 965 --RYGNNDLDKYNNNI--TAVFQNGD-HLPNGQKTROROEFGTYGKDIWKMGALALOE 1019
 176 GLSESKKNNILDYSYNKLN-NAEKDDCCLEKFAKPOPLRWYEVMSDEFCRRKKLEDK 234
 1020 A---GGKTLTETYNSTVFNHGLTGKTLNFAFSRSEFLRMWTEGDDCRIRITQLOI 1076
 235 VEDVICAKADYEGCK-NKNSMNSCVKVEYENYITGKKTQYESQEGKFNTERKROKPEY 293
 1077 LKRRCM-VYQYNDKDGDKDKCECTEACTYTKKMLNMDNYKKQNGRY-TEVKGISP-Y 1133
 294 NYSYKADASEY 304
 DB 1134 KEDSDYKESKY 1144

RESULT 15

097312 PRELIMINARY; PRT; 2169 AA.
 ID 097312
 AC 097312
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DT var (3D7-varT3-2).
 PFC1120C, MAL3P7.55.
 Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3D7;
 RX MEDLINE=99376085; PubMed=10448855;
 RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
 RA Churcher C.M., Craig A., Davies R.M., Devlin K., Felwell T.,
 RA Gentles S., Gwilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
 RA Horrocks P., Jagels K., Jassal B., Kyes S., McLean J., Moule S.,
 RA Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
 RA Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
 RA Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
 RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
 falciparum";
 RL Nature 400:532-538(1999).
 DR EMBL: AL034559; CAB39061.1; -
 DR InterPro: IPR004258; PFEMP.
 DR Pfam: PF03011; PFEMP; 2.
 SO SEQUENCE 2169 AA; 246913 MW; 1D48ACB7AFAE3BE CRC64;

Query Match 23.9%; Score 407.5; DB 5; Length 2169;
 Best Local Similarity 25.8%; Pred. No. 3.1e-21;

Matches 124; Conservative 52; Mismatches 116; Indels 189; Gaps 18;
 QY 10 GCMKPSYDPMDC-----KKNDNSHSGA-CMPRRKQKIC-----VRPLTGGGELTRKPEDITKINCAAKE 70
 DB 850 GCGCKEKPFWKCTISSGSDTGTSTTKONDESGEGHRSKRHTESSDSTTSSGSCVCPVR 909
 39 ROKLCVRLDTQ--GGEIRK-----PED-----ILTKPINCACAKET 71
 DB 910 RRLVYVGLTKMAEFARKSSTSPQPEGVANASASTSSPTDQTLNDATFQSAIET 969
 72 HFAMHKKYKKNV-----AENELSGKRIEGRKQMYTTEGDFRIFGTDIS 119
 DB 970 FFLMHKKYKKNV-----AENELSGKRIEGRKQMYTTEGDFRIFGTDIS 119
 105 TFGDFRIFGTDISSCRYIKDTSQTKSLGDOATTEK-----GDT-----HT 148
 DB 1030 TLGDYRDLGVYKEDYINALKASGDNPNTKLTITQIIEKIKVIEKSGDTPSPRTPGQOP 1089
 149 DNRKLOEWMTIHGPKIWESMLCALN-NGISESEK-----N 184
 DB 1090 SDNDP-KSMWKLAHYIMNGMIYALYDTATAGCKIEKDAVYKKLMDAANKRPKQV 1148
 185 ILDYSYNKLNAEKD-----CCEKFAKPOPLRWYEVMSDEFCRRKKLEDK 234
 DB 1149 GOODTYEKEVEIEEDSGQKASTASQTPSPRASGENKPTLDSFVRRTYFRLYEMGET 1208
 224 FCRERKKLEDKVEDVC-----IAK---DYEGCKNNKSN-----SCV 258
 DB 1209 FCRERKKLEKIVDCEVEENTCARGCTTKOKTSGOEPKSKIVENKDKIFKLEKPSCA 1268
 259 KVCKEYNYITGKKTQYESQEGKFNTERKROKPEYNSYSK-----DASEYL 305
 DB 1269 TPCGLKRWIRKRKDEYKNGKSAVN---EQKTYEKNNGKGGGNGVCGTLDENADFL 1324
 306 K 306
 QY 306 K 306
 DB 1325 K 1325

Search completed: June 20, 2003, 15:05:23
 Job time : 50.1396 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 14:36:44 ; Search time 56.4583 Seconds
(Without alignments)
828.417 Million cell updates/sec

Title: US-10-087-013-9

Perfect score: 1913
Sequence: 1 PCKMVKLISEQIEKNHN.....AKHARDYKLTENMCTNG 351

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Ched: 908470 seqs, 133250620 residues

1.1 number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1913	100.0	351	22 AAB62149	P. falciparum varC
2	646	33.8	793	16 AAR70234	P. falciparum E31a
3	646	33.8	921	18 AAM22480	P. falciparum E31a
4	646	33.8	921	21 AAV77902	P. falciparum ebl-1
5	625.5	32.7	2703	16 AAR70236	P. falciparum Proj
6	625.5	32.7	2710	18 AAM22482	Plasmodium Proj
7	625.5	32.7	2710	21 AAV77904	P. falciparum Proj
8	625.5	32.7	3060	18 AAM22475	Plasmodium var-7
9	625.5	32.7	3060	21 AAV77905	Plasmodium var-7
10	550.5	28.8	3542	22 AAB62142	P. falciparum FC83

11	548	28.6	407	22 AAB62148	P. falciparum varC
12	509.5	26.6	294	22 AAB62147	P. falciparum varC
13	444	23.2	308	22 AAB62151	P. falciparum varC
14	440	23.0	311	22 AAB62150	P. falciparum varC
15	412.5	21.6	2913	17 AAM00384	Plasmodium falcipa
16	398.5	20.8	2197	21 AAB18352	Plasmodium falcipa
17	375	19.6	2228	20 AAM93944	Plasmodium falcipa
18	374.5	19.6	2182	18 AAM22476	Plasmodium var-1
19	374.5	19.6	2182	21 AAV77906	Plasmodium var-1
20	355	18.6	1726	17 AAM00385	Truncated Plasmod
21	353	18.5	749	16 AAR70233	P. falciparum EBL
22	353	18.5	749	18 AAR22479	Plasmodium ebl-1
23	353	18.5	749	21 AAV77901	P. falciparum ebl-1
24	353	18.5	1086	23 AAV76760	Plasmodium falcipa
25	333.5	17.4	1143	23 AAV76759	Plasmodium falcipa
26	329	17.2	350	23 ABB07656	P. falciparum EBA
27	329	17.2	1435	16 AAR70232	P. falciparum SABP
28	329	17.2	1435	18 AAM22477	Silic acid bindin
29	329	17.2	1435	21 AAV77900	P. falciparum SABP
30	329	17.2	1604	16 AAR70105	TNF-R-EBA 175 fusl
31	329	17.2	1786	14 AAR41043	CD4-EBA175 fusion
32	312	16.3	1421	23 AAV76764	Plasmodium falcipa
33	311	16.3	616	23 AAM50533	Unidentified amino
34	303.5	15.9	1700	21 AAB18144	Plasmodium falcipa
35	284	14.8	324	21 AAV77911	Plasmodium DBL gen
36	272.5	14.2	311	21 AAV77915	Plasmodium DBL gen
37	269.5	14.1	1245	16 AAR70106	TNF-R-Pl. vivax Du
38	269	14.1	329	23 ABB07654	P. vivax PvDBP pro
39	269	14.1	1028	14 AAR41044	Plasmodium vivax Pl
40	269	14.1	1061	16 AAR70231	P. vivax DBP. Pl
41	269	14.1	1115	12 AAR13457	Duffy receptor. P
42	269	14.1	1115	18 AAM22478	Duffy antigen bind
43	269	14.1	1115	21 AAV77899	P. vivax DBP bind
44	260	13.6	1807	22 AAB85697	Recombinant protei
45	260	13.6	2028	22 AAB85698	Recombinant protei

ALIGNMENTS

RESULT 1	
AAB62149	
ID AAB62149 standard; peptide: 351 AA.	
AC AAB62149;	
XX 29-MAY-2001 (first entry)	
DT P. falciparum varCSA polypeptide AAtresDBL3-gamma.	
DE FC83, varCSA protein; chondroitin sulfate A; CSA; var gene; PFEMP1;	
XX erythrocyte membrane protein 1; parasitized red blood cell; PRBC;	
KW malaria; protozoacide; AAtresDBL3-gamma.	
KM Plasmodium falciparum.	
XX W0200116326-A2.	
XX W0200116326-A2.	
XX 08-MAR-2001.	
PD 01-SEP-2000; 2000WO-US24195.	
XX 01-SEP-1999; 99US-0152023.	
PR (USSH) US DEPT HEALTH & HUMAN SERVICES.	
PA Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;	
XX Gysin J, Pouvelle B, Fujii N, Smith J;	
DR WPI; 2001-235109/24.	
XX Novel FC83, varCSA protein, useful for modulating parasitized red blood	
PT cell binding, sequestration and onset of maternal malaria -	

XX Claim 54; Page 73-74; 78pp; English.

PS

XX The invention relates to a P. falciparum FC3, varCSA protein, that is

CC capable of binding to Chondroitin sulfate A (CSA). The var gene and the

CC corresponding P.falciparum erythrocyte membrane protein 1 (PEMP1)

CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The

CC protein and the encoding gene are useful for treating and preventing

CC maternal malaria or in a patient afflicted with maternal malaria. The

CC present sequence represents a P. falciparum varCSA polypeptide

CC AddressB13.gamma.

CC

XX Sequence 351 AA;

SQ

Query Match 100.0%; Score 1913; DB 22; Length 351;

Best Local Similarity 100.0%; Pred. No. 3.6e-167;

Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PCKMVKLISEQIEKNNIHNCKKTEDAKMKCENTKLGEDEGVCMPRRONLCVHYTLKLN 60

1 PCKMVKLISEQIEKNNIHNCKKTEDAKMKCENTKLGEDEGVCMPRRONLCVHYTLKLN 60

DB 121 ICDLDTISEKIDHDVTTAKKITAVFQKIGSKTNTKKVLEREGMKKEGSLTWKMLC 180

121 ICDLDTISEKIDHDVTTAKKITAVFQKIGSKTNTKKVLEREGMKKEGSLTWKMLC 180

QY 181 ALSYNETKMKDESVRTYLMKYYIKNNDIKEYLEEFASRPPELFWYEMEGEDFYKRNKE 240

181 ALSYNETKMKDESVRTYLMKYYIKNNDIKEYLEEFASRPPELFWYEMEGEDFYKRNKE 240

DB 181 ALSYNETKMKDESVRTYLMKYYIKNNDIKEYLEEFASRPPELFWYEMEGEDFYKRNKE 240

181 ALSYNETKMKDESVRTYLMKYYIKNNDIKEYLEEFASRPPELFWYEMEGEDFYKRNKE 240

QY 241 LVSLAKKCDSCITLRNNGTSNKTCDNENCGACTQCEKYKMMERKKHYSQKKRFOLY 300

241 LVSLAKKCDSCITLRNNGTSNKTCDNENCGACTQCEKYKMMERKKHYSQKKRFOLY 300

DB 241 LVSLAKKCDSCITLRNNGTSNKTCDNENCGACTQCEKYKMMERKKHYSQKKRFOLY 300

241 LVSLAKKCDSCITLRNNGTSNKTCDNENCGACTQCEKYKMMERKKHYSQKKRFOLY 300

QY 301 KNSATYNGLAVEANSEYTKNDPEVTEANSKHAHRYLTQLENNMCTING 351

301 KNSATYNGLAVEANSEYTKNDPEVTEANSKHAHRYLTQLENNMCTING 351

DB 301 KNSATYNGLAVEANSEYTKNDPEVTEANSKHAHRYLTQLENNMCTING 351

301 KNSATYNGLAVEANSEYTKNDPEVTEANSKHAHRYLTQLENNMCTING 351

RESULT 2

AA70234

ID AAR70234 standard; Protein; 793 AA.

XX

AC AAR70234;

XX

XX 22-SEP-1995 (first entry)

XX

DE P. falciparum E31a.

XX

KW Erythrocyte binding ligand; E31a; binding domain; malaria; therapy;

XX vaccine.

XX Plasmodium falciparum.

OS

PN W09507353-A.

XX

PD 16-MAR-1995.

XX

XX 07-SEP-1994; 94W0-US10230.

PF

XX 10-SEP-1993; 93US-0119677.

PR

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA

XX Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;

PI Wellens TE;

XX WPI; 1995-123427/16.

DR

DR N-PSDB; AA083527.

XX

XX New erythrocyte binding domain polypeptide(s) - isolated from

PT Plasmodium binding proteins, used in diagnosis, treatment and

PT prevention of malaria

XX

PS Disclosure; Page 51-52; 81pp; English.

XX

CC Erythrocyte binding ligand (EBL) family genes were cloned from

CC P. falciparum chromosome 7 subsegment libraries constructed during

CC genetic studies of the chloroquine resistance locus. The 4 genes,

CC EBL-e1 (AA083526), E31a (AA083527), EBL-e2 (AA083528) and ProJ3

CC (AA083529), encode the proteins given in AAR70233-36, respectively. The

CC binding domains of such proteins can be expressed e.g. in E. coli.

CC yeast, mammalian, insect and in vacinia virus and adenovirus-infected

CC cells, and provide protection against P. falciparum.

CC

XX Sequence 793 AA;

SQ

Query Match 33.8%; Score 646; DB 16; Length 793;

Best Local Similarity 41.0%; Pred. No. 2.9e-50;

Matches 150; Conservative 54; Mismatches 112; Indels 50; Gaps 16;

QY 2 CKMVKLISEQIEKNNIHNCKKTED-AKMKCENT--KLGEDEGVCMPRRONLCVHYTLK 58

396 CEIYAEMLKDKNGRTYVGEYRKETSEWTCDSKIKMG-HGACIPRRQKCLHYLEK 454

DB 59 LNDKREDELEAFIKSAAEFTLLROY--NSKVED-DKLLHRDMIPPEFRSMFYT 114

455 IMNTNE--LKVAFIKCAAEFTLLWONYKKDKNGNAEDLEKGLIPEDFKROMFYT 512

QY 115 FGDRDCLDTDI-SEKIDHDVTTAKKITAVFQKIGSKTNTKKVLEREGMKKEGSL 173

513 FADYRDLCLDTDISSKQDTSKYGKVCNIDVFKISN-----SIRYKSWETNGPY 566

DB 174 IMKGMCAISYNT-----ET-KKMDGVRTYLMKYYIKNNDIKEYLEEFASRPPELFW 225

567 IMEGMCAISYDPSLNVNPETHKLLTEGNNE-EKVI-F-GSDSFTLLSFSPQFLRW 624

QY 226 VTEMGDEPFYNNRKKELVSLKKCDSCITLRNNGTSNKTCDNENCGACTQCEKYKMMER 285

625 LTEMGENFECKQKREYKVLAKCKDCDDVDDG---KC--NGKCVACKDQCKOYHSWIGI 678

DB 286 WKKHYSQKKKPOLYKNSATYNNGLAVKEANSEYTKNDPEVTEANSKHAHRYLTQLEN 345

679 WIDNYKKQKGRYEVK-----KILPYKEDKRYKNSD---ARDLYKTQLOLY 721

QY 346 MICTNG 351

DB 722 MKCVNG 727

RESULT 3

AAW22480

ID AAW22480 standard; Protein; 921 AA.

XX

AC AAW22480;

XX

XX 07-OCT-1997 (first entry)

XX

DE Plasmodium E31a.

XX

XX DBL gene family; SAMP; sialic acid binding protein; vaccine; therapy;

KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;

KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;

XX Plasmodium.

OS

PN Plasmodium falciparum.

XX

XX W09640766-A2.

PN

XX 19-DEC-1996.

PD

XX

PF 07-JUN-1996; 96MO-US09508.
 XX
 PR 07-JUN-1995; 95US-0487826.
 XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;
 PI Wellens TE;
 XX

DR WPI: 1997-052231/05.
 DR N-PSDB: AAT72895.
 XX

PT New malaria vaccines - contains cysteine-rich DBL family protein
 PT binding domains homologous domains of the Duffy and sialic acid
 PT binding proteins
 XX

PS Disclosure: Page 43-45; 96pp; English.
 XX

This sequence represents E31a of Plasmodium. E31a belongs to the Duffy binding like (DBL) family of genes which have homology to the Duffy antigen binding protein (DABP) and sialic acid binding protein (SABP) conserved regions (see AAT72889 and AAT72888 respectively). The var family of genes modulate cytoadherence and antigenic variation of Plasmodium infected erythrocytes. SABP and the Duffy antigen binding protein (DABP) are soluble proteins that appear in the culture supernatant after infected erythrocytes release merozoites. DABP and SABP mediate the binding of merozoites and schizonts to the erythrocyte surface. These proteins are necessary for erythrocyte invasion by the parasite. This sequence can be used in the compositions of the invention. The compositions are for the treatment and prevention of malaria, and comprise either a nucleotide sequence or encoded polypeptide of the var-1, var-2, var-3 or var-7 genes of the DBL gene family, a family of genes having homology with conserved regions of DABP and SABP. The compositions are used for the treatment and prevention of malaria. They are also used in the preparation of vaccines for inducing a protective immune response in a mammal to Plasmodium merozoites (especially Plasmodium falciparum or Plasmodium vivax).

SO Sequence 921 AA;

Query Match 33.8%; Score 646; DB 16; Length 921;
 Best Local Similarity 41.0%; Pred. No. 3.6e-50;
 Matches 150; Conservative 54; Mismatches 112; Indels 50; Gaps 16;

2 CKMVQXLSIOIEKNINHNCKTDED-AKMKCENT--KIGDEGVCMPRRONLCVHYLTK 58
 396 CEIYAEMLKDNKNGTTEGECTRKETYSMTDDESKIKMGQ-HGACIPPRKCLHYLEK 454
 59 LINDSKEDDLREAFIKSAAEFTLLROY--NSKNVED-DKILHRDIPPEFRSMFYT 114
 455 IMTWTNE--LKVAETKCAAEFTLLQNYKKDKNGNAMEDDEKLKGGIIPDFRROMFYT 512
 115 FGVDRLDCLDTI-SEKIAHDVTTAKKKTITAVFOKIGSKTTNGKKVLEBGMKKEGELS 173
 513 FADYRDICLGTDISSKDTSKGVKVCNIDVFEYKISN-----SIRYRSMTWETNGPV 566
 174 IMKGMALCALSYNT-----ET-KKMDGVFTYLMKTYKKNDDIKEYLEEFASPPFLRW 225
 567 IMEGMLCALSDYDTSLNWNVPETHKKLTBGNMNF-EKYTF-GSDSSTLTSKSERPQFLRW 624
 226 VTEMGEDEVKNNRKKELVSLKRCDCSTLRNNGTSNKTCDNENCGACTQCEYKKWMER 285
 625 LTEMGEDEVKNNRKKELVSLKRCDCSTLRNNGTSNKTCDNENCGACTQCEYKKWMER 285
 286 WKRIYSQKKKFFOLYKKSATYNNGLAVKANSSEYKNDPEVTEANSKHAHDYLTQLEN 345
 679 WIDYKXKQKGRXYEVK-----KIPLYKEDKDVKNSSD--ARDYLTQLEN 721
 OY 346 MICTNG 351
 DB 722 MKCYNG 727

RESULT 4
 AAT77902
 ID AAT77902 standard; Protein: 921 AA.
 XX

AC AAT77902;
 XX

DT 13-JUN-2000 (first entry)
 XX

DE P. falciparum ebl-1 related polypeptide.
 XX

KW DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
 KW DABP; Sialic Acid Binding Protein; SABP; malaria; vaccine; immunisation;
 KW protozoacide.
 XX

OS Plasmodium falciparum.
 XX

PN US593827-A.
 XX

PD 30-NOV-1999.
 XX

PF 07-JUN-1995; 95US-0487826.
 XX

PR 10-SEP-1993; 93US-0119677.
 XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX

PI Sim KL, Chitnis C, Peterson DS, Su X, Wellens TE, Miller LH;
 XX

DR WPI: 2000-194196/17.
 DR N-PSDB: AA298284.
 XX

PT Isolated protein binding domains from Plasmodium vivax and Plasmodium
 PT falciparum erythrocyte binding proteins useful for vaccinating against
 PT malaria -
 XX

PS Disclosure: Columns 61-66; 93pp; English.
 XX

The invention relates to ebl-1 polypeptides that are encoded by the DBL (Duffy-binding like) gene family. The ebl-1 proteins are substantially identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid Binding Protein (SABP), which are soluble proteins that appear in the culture supernatant after erythrocytes infected with malaria release merozoites. Immunochemical studies indicate that DABP and SABP are the respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be used to vaccinate against malaria, especially caused by P. falciparum. Immunization with the polypeptide provides effective protection against malaria.

SO Sequence 921 AA;

Query Match 33.8%; Score 646; DB 21; Length 921;
 Best Local Similarity 41.0%; Pred. No. 3.6e-50;
 Matches 150; Conservative 54; Mismatches 112; Indels 50; Gaps 16;

2 CKMVQXLSIOIEKNINHNCKTDED-AKMKCENT--KIGDEGVCMPRRONLCVHYLTK 58
 396 CEIYAEMLKDNKNGTTEGECTRKETYSMTDDESKIKMGQ-HGACIPPRKCLHYLEK 454
 59 LINDSKEDDLREAFIKSAAEFTLLROY--NSKNVED-DKILHRDIPPEFRSMFYT 114
 455 IMTWTNE--LKVAETKCAAEFTLLQNYKKDKNGNAMEDDEKLKGGIIPDFRROMFYT 512
 115 FGVDRLDCLDTI-SEKIAHDVTTAKKKTITAVFOKIGSKTTNGKKVLEBGMKKEGELS 173
 513 FADYRDICLGTDISSKDTSKGVKVCNIDVFEYKISN-----SIRYRSMTWETNGPV 566
 174 IMKGMALCALSYNT-----ET-KKMDGVFTYLMKTYKKNDDIKEYLEEFASPPFLRW 225
 567 IMEGMLCALSDYDTSLNWNVPETHKKLTBGNMNF-EKYTF-GSDSSTLTSKSERPQFLRW 624
 226 VTEMGEDEVKNNRKKELVSLKRCDCSTLRNNGTSNKTCDNENCGACTQCEYKKWMER 285

Db 625 LTEMGENECKEOKKEKVKLLAKCKDCDDVDG-----KC--NGKCVACKCKOCKOYHSMWIGI 678

QY 286 WKHHYSOQKKFOLYKNSATYNNGLAVKEANSEYTKNDPEVTEANSAKHARDYLKTOLEN 345

Db 679 WIDNYKKOKRGYEVK-----KIPLYEKDKYKNSD---ARDLKTOLON 721

QY 346 MCTNG 351

Db 722 MKCVNG 727

RESULT 5

AAR70236

ID AAR70236 standard; Protein; 2703 AA.

AC AAR70236;

XX 22-SEP-1995 (first entry)

DT 22-SEP-1995 (first entry)

XX 22-SEP-1995 (first entry)

DE P. falciparum Proj3.

Erythrocyte binding ligand; Proj3; binding domain; malaria; therapy; vaccine.

XX Plasmodium falciparum.

OS WO9507353-A.

XX 16-MAR-1995.

XX 07-SEP-1994; 94WO-US10230.

XX 10-SEP-1993; 93US-0119677.

PR (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;

PI Wellens TE;

XX WPI: 1995-123427/16.

DR N-PSDB; AA083529.

XX New erythrocyte binding domain polypeptide(s) - isolated from

PT Plasmodium binding proteins, used in diagnosis, treatment and

PT prevention of malaria

XX Disclosure; Page 61-65; 81pp; English.

XX Erythrocyte binding ligand (EBL) family genes were cloned from

CC P. falciparum chromosome 7 subsegment libraries constructed during

CC genetic studies of the chloroquine resistance locus. The 4 genes,

CC EBL-e1 (AA083526), E31a (AA083527), EBL-e2 (AA083528) and Proj3

CC (AA083529), encode the proteins given in AAR70233-36, respectively. The

CC binding domains of such proteins can be expressed e.g. in E. coli

CC yeast, mammalian, insect, and in vaccinia virus and adenovirus-infected

CC cells, and provide protection against P. falciparum.

CC

SO Sequence 2703 AA;

QY Query Match 32.7%; Score 625.5; DB 16; Length 2703;

Best Local Similarity 38.8%; Pred. No. 1.3e-47;

Matches 137; Conservative 61; Mismatches 126; Indels 29; Gaps 9;

Db 2 CKMKOKLISDIEKNNHNCCKTD-AKWCENTKLEDEGVCMPPRONLCVHYL---T 57

1342 CKYIEKILEGNGRTTVEGECNPKESYDMDCKNNIDISHDGCMPPROKCLVYIAHES 1401

QY 58 KLNDSEEDREAFITSAAEFTLLROYNSKRVEDDKILHMTPEPFRRSMFTFGD 117

Db 1402 QTEWIKTDNDKDFITAAAEFTLSWQYKSKNDSEAKILDRILISOFLRSMMYTFGD 1461

QY 118 YRDICLDTDISEKADHDVTTAKKRITAVFOKISGKTTNGKLYEREGGWKEVGLSTWKG 177

|||||:|||||: || || || || |||: || || || || |||

Db 1462 YRDICLNTDISKK--QNDVAKAKDKIGKFFESDGSKSPSG---LSQHEWKTNGPEIMKG 1516

QY 178 MICALS-YNTEERKKNDEGRTYLMKIYKNNDIKETLEEFASRPPELRVWTEGDEPYKN 236

Db 1517 MICALTKRYTDTDNKRKINDSYDKVNSQNSQNGSLSEFAKPOFLRMIMIEGEEFCAE 1576

QY 237 RKKEIYSLKKKDCSTLRNNGTSNKRQDNNENCGACKTCQEKYKKMKEMKKHHYSOQKK 296

Db 1577 ROKKENIYIKDAC-----NEINSTOOCNDAKH--RCNACRAVQYEVKRRKEFGQTN 1628

QY 297 FOLYKNSATYNNGLAVKEANSEYTKNDPEVTEANSAKHARDYLKTOLENMICT 349

Db 1629 FVLKANVQPD-----PEYKGYEKDGVPQIGN-----EVLKCKCNKKKS 1670

RESULT 6

AAW22482

ID AAW22482 standard; Protein; 2710 AA.

XX AAW22482;

XX 07-OCT-1997 (first entry)

DT 07-OCT-1997 (first entry)

XX 07-OCT-1997 (first entry)

DE Plasmodium Proj3.

XX DBL gene family; SABP; static acid binding protein; vaccine; therapy;

KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;

KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; Immune response;

XX Plasmodium.

XX Plasmodium falciparum.

OS WO9640766-A2.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US09508.

XX 07-JUN-1995; 95US-0487826.

PR (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;

PI Wellens TE;

XX WPI: 1997-052231/05.

DR N-PSDB; AAT72897.

XX New malaria vaccines - contains cysteine-rich DBL family protein

PT binding domains homologous domains of the Duffy and static acid

PT binding proteins

XX Disclosure; Page 50-56; 96pp; English.

XX This sequence represents Proj3 of Plasmodium. Proj3 belongs to

CC the Duffy binding like (DBL) family of genes which have homology to the

CC Duffy antigen binding protein (DABP) and static acid binding protein

CC (SABP) conserved regions (see AAT72889 and AAT72888 respectively). The

CC var family of genes modulate cytoadherence and antigenic variation of

CC Plasmodium infected erythrocytes. SABP and the Duffy antigen binding

CC protein (DABP) are soluble proteins that appear in the culture

CC supernatant after infected erythrocytes release merozoites. DABP and

CC SABP mediate the binding of merozoites and schizonts to the erythrocyte

CC surface. These proteins are necessary for erythrocyte invasion by the

CC parasite. This sequence can be used in the compositions of the

CC invention. The compositions are for the treatment and prevention of

CC malaria, and comprise either a nucleotide sequence or encoded polypeptide

CC of the var-1, var-2, var-3 or var-7 genes of the DBL gene family. a

CC family of genes having homology with conserved regions of DABP and SABP.

CC The compositions are used for the treatment and prevention of malaria.

CC They are also used in the preparation of vaccines and prevention of

CC protective immune response in a mammal to Plasmodium merozoites

CC (especially Plasmodium falciparum or Plasmodium vivax).

XX Sequence 2710 AA: 32.7% Score 625.5; DB 18; Length 2710;
 Query Match Best Local Similarity 38.8%; Pred. No. 1.3e-47;
 Matches 137; Conservative 61; Mismatches 126; Indels 29; Gaps 9;

2 CKWVKLISQIIEKNNHNCCKTED-AKKCENTKLGEEGCVMPPRRNLGVHVL---T 57
 1349 CKIVEKILBGNKGRFTVGCNPKRESYPDMCKNNIDISHDGAQMPPRKLCLYIAHES 1408

QY 58 KLNDSEKEDLRFAFKSAAEFTLLROYNSKNVEDDKILHDMIPPEFRSMFTFGD 117
 1409 QENIKITDNLKDAFIKTAAEFTLSWQYKSKNDESAKILDRGLIPSOFLRSMYTFGD 1468

QY 118 YRDICLDTISEKIAHDVTTAKKITAFAVOKIGSKTNGKVKLEREGMKKEYGLSTWG 177
 1469 YRDICLNTDISKR--ONDAKAKADKIGKFFSKDGSKSPSG---LSRQEMWKTGPEIWK 1523

DB 178 MCALSL-YNTEFKKMDGVTYLMKYIKKNDIKELYEFASRPFLRWTEWGEDFVK 236
 1524 MCALFKRYVTDNKRRIKNDYSYDKVNSQNGNPISLEEFAPKPFRLRMWIEGEFCAL 1583

QY 237 RKRELYSLKKKCDSCFLRNNGTSNKTCDNENCGACKTOCEYKKMKMERKKAHYSQKK 296
 1584 ROKKENIIRKDAC-----NEINSTQOCNDAKH--RCNOACRAVQEVYENKKEFSGQTN 1635

QY 297 FOLYKNSATYNNGLAVKEANSEYKNDPEYTEANSAKHARDYKLTQLENNICT 349
 1636 FVLKANVQPOD-----PEYKGYEKDGVQPIQGN-----EYLQKCDNNKCS 1677

DB 1636 FVLKANVQPOD-----PEYKGYEKDGVQPIQGN-----EYLQKCDNNKCS 1677

RESULT 7
 AAY77904
 ID AAY77904 standard; protein; 2710 AA.
 AC AAY77904;
 DT 13-JUN-2000 (first entry)

DE P. falciparum Proj3 binding domain polypeptide.
 XX DBL gene; Duffy-binding like gene; ebl-1; Duffy Antigen Binding Protein;
 KW DABP; Sialic Acid Binding Protein; SAMP; malaria; vaccine; immunisation;
 KM protozoicide; Proj3.
 XX Plasmodium falciparum.

US5993827-A.
 PD 30-NOV-1999.
 XX 07-JUN-1995; 95US-0487826.
 PF 10-SEP-1993; 93US-0119677.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Sim KL, Chittis C, Peterson DS, Su X, Wellens TE, Miller LH;
 XX WPI; 2000-194198/17.
 DR N-PSDB; AA298286.
 XX Isolated protein binding domains from Plasmodium vivax and Plasmodium
 PT falciparum erythrocyte binding proteins useful for vaccinating against
 PT malaria -
 PS Disclosure; Columns 79-92; 93pp; English.
 XX The invention relates to ebl-1 polypeptides that are encoded by the DBL
 CC (Duffy-binding like) gene family. The ebl-1 proteins are substantially
 CC identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid
 CC Binding Protein (SABP), which are soluble proteins that appear in the

CC culture supernatant after erythrocytes infected with malaria release
 CC merozoites. Immunochemical studies indicate that DABP and SABP are the
 CC respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy
 CC and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be
 CC used to vaccinate against malaria, especially caused by P. falciparum.
 CC Immunization with the polypeptide provides effective protection against
 CC malaria. The present sequence represents the Proj3 binding domain
 CC polypeptide.

XX Sequence 2710 AA: 32.7% Score 625.5; DB 21; Length 2710;
 Query Match Best Local Similarity 38.8%; Pred. No. 1.3e-47;
 Matches 137; Conservative 61; Mismatches 126; Indels 29; Gaps 9;

2 CKWVKLISQIIEKNNHNCCKTED-AKKCENTKLGEEGCVMPPRRNLGVHVL---T 57
 1349 CKIVEKILBGNKGRFTVGCNPKRESYPDMCKNNIDISHDGAQMPPRKLCLYIAHES 1408

QY 58 KLNDSEKEDLRFAFKSAAEFTLLROYNSKNVEDDKILHDMIPPEFRSMFTFGD 117
 1409 QENIKITDNLKDAFIKTAAEFTLSWQYKSKNDESAKILDRGLIPSOFLRSMYTFGD 1468

QY 118 YRDICLDTISEKIAHDVTTAKKITAFAVOKIGSKTNGKVKLEREGMKKEYGLSTWG 177
 1469 YRDICLNTDISKR--ONDAKAKADKIGKFFSKDGSKSPSG---LSRQEMWKTGPEIWK 1523

DB 178 MCALSL-YNTEFKKMDGVTYLMKYIKKNDIKELYEFASRPFLRWTEWGEDFVK 236
 1524 MCALFKRYVTDNKRRIKNDYSYDKVNSQNGNPISLEEFAPKPFRLRMWIEGEFCAL 1583

QY 237 RKRELYSLKKKCDSCFLRNNGTSNKTCDNENCGACKTOCEYKKMKMERKKAHYSQKK 296
 1584 ROKKENIIRKDAC-----NEINSTQOCNDAKH--RCNOACRAVQEVYENKKEFSGQTN 1635

QY 297 FOLYKNSATYNNGLAVKEANSEYKNDPEYTEANSAKHARDYKLTQLENNICT 349
 1636 FVLKANVQPOD-----PEYKGYEKDGVQPIQGN-----EYLQKCDNNKCS 1677

DB 1636 FVLKANVQPOD-----PEYKGYEKDGVQPIQGN-----EYLQKCDNNKCS 1677

RESULT 8
 AAW22475
 ID AAW22475 standard; protein; 3060 AA.
 AC AAW22475;
 DT 12-SEP-1997 (first entry)

DE Plasmodium var-7.
 XX DBL gene family; SABP; sialic acid binding protein; vaccine; therapy;
 KW Duffy binding like gene; Duffy antigen binding protein; erythrocyte;
 KW DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;
 KM Plasmodium.
 XX Plasmodium vivax.
 OS Plasmodium falciparum.
 OS Plasmodium var-7.
 PN W09640766-A2.
 XX 19-DEC-1996.
 PD 07-JUN-1996; 96WO-US09508.
 PF 07-JUN-1995; 95US-0487826.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Chittis C, Miller LH, Peterson DS, Sim KL, Su X;
 XX Wellens TE;
 PI WPI; 1997-052231/05.
 DR N-PSDB; AAT72882.

XX New malaria vaccines - contains cysteine-rich DBL family proteins
PT binding domains homologous domains of the Duffy and Xk-like
PT binding proteins
XX
PS Claim 8; Page 61-67; 96pp; English.

Claim 8; Page 61-67; 96pp; English.

CC This sequence represents var-7 of Plasmodium. Var-7 belongs to
CC the Duffy binding like (DBL) family of genes which have homology to the
CC Duffy antigen binding protein (DABP) and sialic acid binding protein
CC (SABP) conserved regions (see AA72869 and AA72888 respectively). The
CC var family of genes modulate cytoadherence and antigenic variation of
CC Plasmodium infected erythrocytes. SABP and the Duffy antigen binding
CC protein (DABP) are soluble proteins that appear in the culture
CC supernatant after infected erythrocytes release merozoites. DABP and
CC SABP mediate the binding of merozoites and schizonts to the erythrocyte
CC surface. These proteins are necessary for erythrocyte invasion by the
CC parasite. This sequence can be used in the compositions of the
CC invention. The compositions are for the treatment and prevention of
CC malaria, and comprise either a nucleotide sequence or encoded polypeptide
CC of the var-1, var-2, var-3 or var-7 genes of the DBL gene family. A
CC family of genes having homology with conserved regions of DABP and SABP.
CC The compositions are used for the treatment and prevention of malaria.
CC They are also used in the preparation of vaccines for inducing a
CC protective immune response in a mammal to Plasmodium merozoites
CC (especially Plasmodium falciparum or Plasmodium vivax).

SQ Sequence 3060 AA;

Query March	32.7%	Score 625.5;	DB 18;	Length 3060;
Best Local Similarity	38.8%	Pred. No. 1.5e-47;		
Matches 137; Conservative	61;	Mismatches 126;	Indels 29;	Gaps 9.

[illegible]

RESULT 9

AAV77905 standard; Protein; 3060 AA.

AC	AA77905;
XX	
DT	13-JUN-2000 (first entry)
XX	

XX. Plasmodium var-1 polypeptide.

DABP: Sialic Acid Binding Protein

Plasmodium sp.
XX
KM
XX
protozoacide; var-7.
SADP; malidid; vaccine; immunisation;

Plasmodium sp.

Plasmodium sp.

XX US5993827-A.
PN

PD 30-NOV-1999.

07-JUN-1995; 95US-0487826.

PR 10-SEP-1993; 9305-0119677.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PL SLM KL, Chitnis C, Peterson DS, Su X, Willems TE, Miller LH, .xy

DK WPL; 2000-194198/17.
DB M-PDD; 19700007

XX
DE

PT *Calcioparum erythrocyte binding proteins useful for vaccinating against*

PS Disclosure; Columns 109-124; 93pp; English.

The invention relates to eb1-1 polypeptides that are encoded by the DB1 (Duffy-binding like) gene family. The eb1-1 proteins are substantially identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid Binding Protein (SABP), which are soluble proteins that appear in the culture supernatant after erythrocytes infected with malaria release merozoites. Immunochemical studies indicate that DABP and SABP are the respective ligands for Plasmodium vivax and Plasmodium falciparum Duffy and sialic acid receptors on erythrocytes. The eb1-1 polypeptides may be used to vaccinate against malaria, especially caused by P. falciparum. Immunization with the polypeptide provides effective protection against malaria. The present sequence represents the var-7 polypeptide.

50 Sequence 3060 AA;

Query Match	32.78;	Score 625.5;	DB 21;	Length 3060;
Best Local Similarity	38.88;	Pred. No. 1.5e-47;		
Matches 137; Conservative	61;	Mismatches 126;	Indels 29;	Gaps 9

```

QY 2CMMVKLISEQLEKNNNNHNCCKTDE-AKMKENKRLGEDGVGMPPROULCYHYH---T 57
Db 1347 CKIVERKLEGGKNGRTTVEGCPNPKRESYPMDCKNNIDISHGACMPRRKOLCLITYAHMS 1407
QY .58 KLNDKSCREEDLEALATIKSAAEITFLRLROYNSKAVVEDKLLIHRMDIPEFFRSMFYFGD 117
Db 1407 QDEENIKTDNDKDAIKTAAAEITFLSMQYKSKNDSFAKLLIDGLIPSLFRSMYTFED 1466
QY 118 YRDICLDITDISEKIDHDVYTAACKRIAYFQIGSKRTNGKRVLEBGGWKEJGSLTWKG 177
Db 1467 YRDICLINTDISK--QNDVAAKAKDIKGFPSKOSKSPSG---LSQDEWMTKNGPRTWKG 1522
QY 178 MCAALS-INTENKKKKDECVRTYLLAKTYIKNNDIKEYLEEFASRPPLRLVWTEGDEYKN 236
Db 1522 MCAALTKLVYTDIDNKRKTKINDYSYDKVNOSONGNSPLEEFAAPQFLRMIMIGGEFFCAE 1581
QY 237 RKKEVLVSUKKCDSCITLNNNGTSNKYTDDBENCGACQTQCEKKRKMMEWKHYSSOKAK 296
Db 1582 ROKKENIITKDC-----NEINSTQOCNDAKH--RCNQACRAQOEVEVKKKKHFEFSQITN 1633
QY 297 POLYNSATYNNGLAVLEANSSETYKNDPEVTEVANSKARHADVLTQLENNMISTC 349
Db 1634 EVLAKANVOPD-----PEYKGYEKKDDVQIOGN-----EYLLKCDNNKS 1675

```

RESULT 10

ID AAB62142 standard; Protein; 3542 AA.

AC AAB62142;

29-MAY-2001 (first entry)

29-MAY-2001 (first entry)

DE P. falciparum FCRI3.varCSA protein.
 XX FCRI3.varCSA protein; chondroitin sulfate A (CSA); var gene; PFEMP1;
 XX erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
 KM malaria; protozoacide.
 OS Plasmodium falciparum.
 XX WO200116326-A2.
 PN
 XX
 PD 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000WO-US24195.
 XX
 PR 01-SEP-1999; 99US-0152023.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
 Gysin J, Pouvelle B, Fujii N, Smith J;
 WPI: 2001-235109/24.
 DR N-PSDB; AAF57301.
 XX
 PT Novel FCRI3.varCSA protein, useful for modulating parasitized red blood
 cell binding, sequestration and onset of maternal malaria -
 XX
 PS Claim 12: Page 63-71; 78pp; English.
 XX
 CC The invention relates to a P. falciparum FCRI3.varCSA protein, that is
 capable of binding to chondroitin sulfate A (CSA). The var gene and the
 CC corresponding P.falciparum erythrocyte membrane protein 1 (PFEMP1)
 CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
 CC protein and the encoding gene are useful for treating and preventing
 CC maternal malaria in a patient afflicted at a risk for contracting
 CC maternal malaria or in a patient afflicted with maternal malaria. The
 CC present sequence represents the P. falciparum FCRI3.varCSA protein.
 CC
 SQ Sequence 3542 AA:
 Query Match 28.8%; Score 550.5; DB 22; Length 3542;
 Best Local Similarity 39.0%; Pred. No. 1.5e-40;
 Matches 133; Conservative 48; Mismatches 115; Indels 45; Gaps 14;
 2 CKWVKLSIQLEKKNINHC--KKTEDA--KKKCENTKGEDEVCMPPRRQKLCVHYLT 57
 1261 CKIVNDILKENDGCKKQVEDCHPKKNSNGIPDMQGINLYEDPRVCMPPRQKLCVHYLT 1320
 58 KLNDGSK---EEDLRFAFTKSAAEFTLLROYNSKNVED--DKILHRDMPPEFFRSW 111
 1321 NDREIKKLSQVNLKFAFTKSAAEFTFSWYKSKDSEGNELDKELKEKIPAPFLRSW 1380
 112 FYTFGGYRDLCDTDSKADHDVTTAKKITAIVQKIGSKTTNGKKVLEBEGMKKEYG 171
 1381 FYTFGGYRDLCDTDSKADHDVTTAKKITAIVQKIGSKTTNGKKVLEBEGMKKEYG 1435
 172 LSIWGMICALSYNTEKMKDEGVRTYLMKVIYKNN--NDIKKEYLEEFASRPFLRWVTE 228
 1436 HEIWEAMLCAL-VKIAKKDD-----FTENYGINNVKFSKSTLLEFARPPFLRWLTE 1489
 229 WGEDFYVKNRKKELVLSKKKCDSCITLRNNGTSNKTCDNENCGA-CKTQCEKVKKMERWK 287
 1490 WYDDYCYTRQKYLKVDQEKCS-----NDQLKCDPECKKKCEDYKYNKK-K 1535
 288 KHYSSOKKKFQLYKNSATYNN---GLAVKE---ANSETYKN 322
 1536 KEMIPQDKYKXDERDKKRFDRQHIGVAVTDYTGTMATDLYN 1576
 RESULT 11
 AAB62148
 ID AAB62148 standard; peptide; 407 AA.
 XX

AC AAB62148;
 XX
 DT 29-MAY-2001 (first entry)
 XX
 DE P. falciparum varCSA polypeptide A4 DBL4-gamma.
 XX
 XX FCRI3.varCSA protein; chondroitin sulfate A (CSA); var gene; PFEMP1;
 KM erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
 KM malaria; protozoacide; A4 DBL4-gamma.
 OS Plasmodium falciparum.
 XX WO200116326-A2.
 PN
 XX
 PD 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000WO-US24195.
 XX
 PR 01-SEP-1999; 99US-0152023.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheidig C;
 Gysin J, Pouvelle B, Fujii N, Smith J;
 WPI: 2001-235109/24.
 DR
 XX
 PT Novel FCRI3.varCSA protein, useful for modulating parasitized red blood
 cell binding, sequestration and onset of maternal malaria -
 XX
 PS Disclosure: Page 72-73; 78pp; English.
 XX
 CC The invention relates to a P. falciparum FCRI3.varCSA protein, that is
 capable of binding to chondroitin sulfate A (CSA). The var gene and the
 CC corresponding P.falciparum erythrocyte membrane protein 1 (PFEMP1)
 CC modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
 CC protein and the encoding gene are useful for treating and preventing
 CC maternal malaria in a patient afflicted at a risk for contracting
 CC maternal malaria or in a patient afflicted with maternal malaria. The
 CC present sequence represents a P. falciparum varCSA polypeptide
 CC A4 DBL4-gamma.
 CC
 SQ Sequence 407 AA:
 Query Match 28.6%; Score 548; DB 22; Length 407;
 Best Local Similarity 34.6%; Pred. No. 1.1e-41;
 Matches 126; Conservative 52; Mismatches 128; Indels 58; Gaps 12;
 10 SEQIER-NNINHCKTEDAKMKCENTKLGEDVCMPPRRQKLCVHYLTKLNDGKEEDL 68
 6 TKKIEECNTRKYPTKNDYPMWNTDQVINEBSCMPPRRQKLCVHYLTKLNDGKEEDL 65
 69 REAFITSAAEFTLLROYNSK-----NVEDD-----KILHRDMPPEFFRSW 113
 66 RKAFICCAIAETFLWMDKYEKKDEKKTGEGISDPPDQPKKLGCTIPEDFKQMFY 125
 114 TFGDYRDLCDTDSKADHDVTTA-KKITAIVQKIGSKTTNGKKVLEBEGMKKEYG 172
 126 TYGDYRDLCDTDSKADHDVTTA-KKITAIVQKIGSKTTNGKKVLEBEGMKKEYG 179
 173 SIWGMICALSYNTEKMKDEGVRTYLMKVIYKNN--DIKEYLEEFASRPFLRWVTE 229
 180 DIWGMICALSYNTEKMKDEGVRTYLMKVIYKNN--DIKEYLEEFASRPFLRWVTE 234
 230 GEDFYVKNRKKELVLSKKKCD--SCTLRNNGTSNKTCDNENCGACCTQCEKVKKMERWK 287
 235 GDQFCRERVYKINQLTGCMEYECGSENGEKKK-----ACKNACCAVYKSWIKDKW 284
 288 KHYSSOKKKFQLYKNSATYNNGLAVKEANSETYKNDEPVTANSARHARDYLKTOLENNI 347
 285 DQEQQAKFD-----KDKKDKKFDGTSAEVVAVAASSVHEVLDQELKN-L 329
 348 CTNG 351
 OY


```

Db      183  -----KNILQDYSYNKLNNAEKDDCDELEFASKPOFLRMVYEMSDSEFCERRK-----LED 233
QY      247  KCDSCITLRNNGTSNKTCDNENCGACKTOCEKYKMMERKKHYSOQKKFOLYKNSATY 306
Db      234  KVEDVICR--AADEYECCKNNKSNNSCVKCKEYENITGKKTQYESQEGKF-----282
QY      307  NNGLAVKEANSEYKNDPEVTEANSAKHARDYLK 340
Db      283  -----NTEKROKKPEY-NSYSKKDASEYIK 306

RESULT 14
AAB62150
ID  AAB62150 standard; peptide; 311 AA.
AC  AAB62150;
XX
XX  29-MAY-2001 (first entry)
XX
XX  P. falciparum varCSA polypeptide FC03 var3DBL-gamma.
XX
XX  FC03, varCSA protein; chondroitin sulfate A; CSA; var gene; PPEMP1;
XX  erythrocyte membrane protein 1; parasitized red blood cell; PRBC;
XX  malaria; protozoacide; FC03 var3DBL-gamma.
XX
XX  Plasmodium falciparum.
XX
XX  WO200116326-A2.
XX
XX  08-MAR-2001.
XX
XX  01-SEP-2000; 2000MO-US24195.
XX
XX  01-SEP-1999; 99US-0152023.
XX
XX  (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX  Scherf A, Miller LH, Gamain B, Baruch DI, Buffet P, Scheldig C;
XX  Gysin J, Pouvelle B, Fujil N, Smith J;
XX  WPI; 2001-235109/24.
XX
XX  Novel FC03, varCSA protein, useful for modulating parasitized red blood
XX  cell binding, sequestration and onset of maternal malaria -
XX
XX  Disclosure; Page 74-75; 78pp; English.
XX
XX  The invention relates to a P. falciparum FC03, varCSA protein, that is
XX  capable of binding to chondroitin sulfate A (CSA). The var gene and the
XX  corresponding P.falciparum erythrocyte membrane protein 1 (PPEMP1)
XX  modulate adhesion of parasitized red blood cell (PRBC) to CSA. The
XX  protein and the encoding gene are useful for treating and preventing
XX  maternal malaria in a patient afflicted at a risk for contracting
XX  CC maternal malaria or in a patient afflicted with maternal malaria. The
XX  CC present sequence represents a P. falciparum varCSA polypeptide
XX  CC FC03 var3DBL-gamma.
XX
XX  Sequence 311 AA:
XX
XX  Query Match 23.0%; Score 440; DB 22; Length 311;
XX  Best Local Similarity 31.9%; Pred. No. 6,3e-32;
XX  Matches 109; Conservative 52; Mismatches 133; Indels 48; Gaps 9;
QY      14  EKNINHC-----KTEDAKKCENTKAGEDEGVCMPRRONLCVHYLTKNDSKEED 67
Db      1  EPTDIDGCMQKXKAGDKYPMDCNSQIHTTHNGACMPRRQKLCVSGLTTRIRAIEX 60
QY      68  LREAFIKSAAEFFELROYNSKNVEDDKTLHRDMTPPEFFRSMYTFPGDYRICLDTPI 127
Db      61  IRTEFIKSAALIEHFMDRKEDNGAELEKNGNIPEGFKRMQYTFGDIYRIFIGRDI 120
QY      128  S-----EKIADHDVTTAKKKTAVFOKISKTTNGKVLREGEQWKEYGSLIWGMICAL 182

```

```

Db      121  SHHAYISGSPRVITILEKENDA---KYAKQNSNNELL--DDMWDQHOKDIMEGLCAL 175
QY      183  SYNTETFKMDEGVRYLYMKYIK--NNDIK--EYLEEPASRPFLRWTEGDEYVKNR 238
Db      176  THKISDEEKKEIKN--KYSYKKILNESPKGSNKVEDFAKKPOFLRMFTMGDEFCQRE 232
QY      239  KELVSLKRRKCDSCITLRNNGTSNKTCDNENCGACKTOCEKYKMMERKKHYSOQKKFQ 298
Db      233  EKEAKKVSC-----SDAKDYDCCKNTKSNASCVSACKVYEDYLTKKRVEYTKQKQFD 286
QY      299  LYKNSATYNNGLAVKEANSEYKNDPEVTEANSAKHARDYLK 340
Db      287  AEKIT-----DKEGYEGESTKDASEYIK 309

RESULT 15
AAM00384
ID  AAM00384 standard; Protein; 2913 AA.
AC  AAM00384;
XX
XX  21-FEB-1997 (first entry)
XX
XX  Plasmodium falciparum erythrocyte membrane protein.
XX
XX  Plasmodium falciparum; erythrocyte membrane protein; malaria;
XX  detection; identification; treatment; prevention; parasite.
XX
XX  Plasmodium falciparum MC type.
XX
XX  OS
XX  FH
XX  Key Location/Qualifiers
XX  Domain 62..394
XX  FT /label= Duffy binding ligand domain 1
XX  FT 607..648
XX  FT /note= "Cysteine rich motif"
XX  FT 839..1272
XX  FT /label= Duffy binding ligand domain 2
XX  FT 1482..1527
XX  FT /note= "Cysteine rich motif"
XX  FT 1706..2005
XX  FT /label= Duffy binding ligand 3
XX  FT 2102..2349
XX  FT /label= Duffy binding ligand 4
XX  FT 2354..2398
XX  FT /note= "Cysteine rich motif"
XX  FT 2450..2475
XX  FT Domain /note= "putative transmembrane domain"
XX
XX  PN MO9633736-A1.
XX
XX  PD 31-OCT-1996.
XX
XX  PF 26-APR-1996; 96MO-US05798.
XX
XX  PR 27-APR-1995; 95US-0430908.
XX
XX  PA (AFTY-) AFTYMAX TECHNOLOGIES NV.
XX
XX  PI Baruch DI, Howard RJ, Pasloske BL;
XX  WPI; 1996-497376/49.
XX  DR N-PSDB; AAT41852.
XX
XX  New Plasmodium falciparum erythrocyte membrane proteins - used to
XX  develop products for the diagnosis, treatment or prevention of
XX  malaria parasite infections
XX
XX  PS Claim 1; Figure 12; 149pp; English.
XX
XX  A polypeptide comprising a Plasmodium falciparum (Pf) erythrocyte
XX  membrane protein 1 (PPEMP1) or active fragments or analogues of that
XX  protein can be used in the treatment or prevention of symptoms of a

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:01:19 ; Search time 19.1745 Seconds
(Without alignments)
538.603 Million cell updates/sec

Title: US-10-087-013-9
Perfect score: 1913
Sequence: 1 PCKMVKLISEQIEKNNIHN.....AKHARDYLKQLENMICTNG 351

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 262574
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	646	33.8	921	2	US-08-568-459A-8
2	646	33.8	921	2	US-08-487-826B-8
3	646	33.8	921	4	US-09-210-288-8
4	625.5	32.7	2710	2	US-08-568-459A-12
5	625.5	32.7	2710	2	US-08-487-826B-12
6	625.5	32.7	2710	4	US-09-210-288-12
7	625.5	32.7	3060	2	US-08-487-826B-14
8	374.5	19.6	2182	2	US-08-487-826B-16
9	353	18.5	749	2	US-08-568-459A-6
10	353	18.5	749	2	US-08-487-826B-6
11	353	18.5	749	4	US-09-210-288-6
12	329	17.2	1435	2	US-08-568-459A-4
13	329	17.2	1435	2	US-08-487-826B-4
14	329	17.2	1435	4	US-09-210-288-4
15	284	14.8	324	2	US-08-568-459A-17
16	284	14.8	324	2	US-08-487-826B-29
17	284	14.8	324	4	US-09-210-288-17
18	272.5	14.2	311	2	US-08-568-459A-21
19	272.5	14.2	311	2	US-08-487-826B-33
20	272.5	14.2	311	4	US-09-210-288-21
21	269	14.1	1115	2	US-08-568-459A-2
22	269	14.1	1115	2	US-08-487-826B-2
23	269	14.1	1115	4	US-09-210-288-2
24	231	12.1	700	2	US-08-568-459A-10
25	231	12.1	700	2	US-08-487-826B-10
26	231	12.1	700	2	US-08-487-826B-10
27	231	12.1	700	4	US-09-210-288-10

28	171.5	9.0	778	6	5198347-4	Patent No. 5198347
29	171	8.9	277	2	US-08-568-459A-15	Sequence 15, Appl
30	171	8.9	277	2	US-08-487-826B-27	Sequence 27, Appl
31	171	8.9	277	4	US-09-210-288-15	Sequence 15, Appl
32	167	8.7	197	6	5198347-2	Patent No. 5198347
33	144.5	7.6	411	2	US-08-568-459A-20	Sequence 20, Appl
34	144.5	7.6	411	4	US-08-487-826B-32	Sequence 32, Appl
35	144.5	7.6	411	4	US-09-210-288-20	Sequence 20, Appl
36	137	7.2	282	2	US-08-568-459A-16	Sequence 16, Appl
37	137	7.2	282	2	US-08-487-826B-28	Sequence 28, Appl
38	137	7.2	282	4	US-09-210-288-16	Sequence 16, Appl
39	135	7.1	291	2	US-08-568-459A-13	Sequence 13, Appl
40	135	7.1	291	2	US-08-487-826B-25	Sequence 25, Appl
41	135	7.1	291	4	US-09-210-288-13	Sequence 13, Appl
42	131.5	6.9	362	2	US-08-568-459A-18	Sequence 18, Appl
43	131.5	6.9	362	2	US-08-487-826B-30	Sequence 30, Appl
44	131.5	6.9	362	4	US-09-210-288-18	Sequence 18, Appl
45	107	5.6	1507	3	US-08-929-329-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-568-459A-8
: Sequence 8, Application US/08568459A
: Patent No. 5849306
: GENERAL INFORMATION:
: APPLICANT: Sim, Kim L.
: APPLICANT: Chitnis, Chetan
: APPLICANT: Miller, Louis H.
: APPLICANT: Peterson, David S.
: APPLICANT: Su, Xinzhan
: TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
: TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Knobe Martens Olson & Bear
: STREET: 620 Newport Center Drive 16th Floor
: CITY: Newport Beach
: STATE: California
: COUNTRY: US
: ZIP: 92660
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/568,459A
: FILING DATE: 07-DEC-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Israel, Ned
: REGISTRATION NUMBER: 29,655
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 235-8550
: TELEFAX: (619) 235-0176
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 921 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Plasmodium falciparum
: US-08-568-459A-8
Query Match 33.8%, Score 646, DB 2, Length 921;

Best Local Similarity 41.0%; Pred. No. 2.4e-50;
Matches 150; Conservative 54; Mismatches 112; Indels 50; Gaps 16;

```

QY 2 CKAWOGLISEQLEKNNHCKKTED-AKMKCENT--KIGEDGVCMPRRONLCVHYLTK 58
Db 396 CEIVAMLMKDKNGRTTVEGCEYKREYSEMTCDSEKTKMQ-HGACIPPRKCLCHLYLK 454
QY 59 LNDKSEEDLRPAFTKSAAEFTFLROY--NSKNVED-DKILHRDMPPEFFSMFYT 114
Db 455 IMTNTNE--LKYAFKCAAEFTFLMÖNYKKDKNGMAEDLEKLGKIIPEDFKRMFT 512
QY 115 PGDYRDICLDPI-SEKIDHDVTTAKKKTITAVFOKISKTTNGKYLREGMKKEYGLS 173
Db 513 FADYRDICLGTDISKKTDSKGVGKVCNIDVFKISN-----SIRYKSMWETNGPV 566
QY 174 IWKGMICALSYNT-----ET-KKMDGVRITYLMKYIKNNDIKEYLEEFASRPFLRW 225
Db 567 IWEGMICALSYDTSLNNVNPETHKKLTGEGNNNF-EKVIF-GSDSTSLSKSEKPOFLRW 624
QY 226 VTEMGEDPVKRNKKEVLSLKKKCDSCITLRNNGTSNKTCDNENCGACKTQCEKRYKKMER 285
Db 625 LTEWGENFCKEÖKREKYLAKCKDODVDGDG---KC--NGKCVACKDCKÖYHSWIGI 678
QY 286 WKHYSSQKKFKPOLYKNSATYNNGLAVKANSEYTKNDPEYFANSKHAADYLTQLEN 345
Db 679 WIDNKKÖKGRYTEVK-----KIPLYEKDVKNSDD--ARDYLTQLEN 721
QY 346 MICTNG 351
Db 722 MKCVNG 727

```

RESULT 2

US-08-487-826B-8
Sequence 8, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B

FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Istreisen, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 921 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-487-826B-8

Query Match

Best Local Similarity 33.8%; Score 646; DB 2; Length 921;
Matches 150; Conservative 54; Mismatches 112; Indels 50; Gaps 16;

```

QY 2 CKAWOGLISEQLEKNNHCKKTED-AKMKCENT--KIGEDGVCMPRRONLCVHYLTK 58
Db 396 CEIVAMLMKDKNGRTTVEGCEYKREYSEMTCDSEKTKMQ-HGACIPPRKCLCHLYLK 454
QY 59 LNDKSEEDLRPAFTKSAAEFTFLROY--NSKNVED-DKILHRDMPPEFFSMFYT 114
Db 455 IMTNTNE--LKYAFKCAAEFTFLMÖNYKKDKNGMAEDLEKLGKIIPEDFKRMFT 512
QY 115 PGDYRDICLDPI-SEKIDHDVTTAKKKTITAVFOKISKTTNGKYLREGMKKEYGLS 173
Db 513 FADYRDICLGTDISKKTDSKGVGKVCNIDVFKISN-----SIRYKSMWETNGPV 566
QY 174 IWKGMICALSYNT-----ET-KKMDGVRITYLMKYIKNNDIKEYLEEFASRPFLRW 225
Db 567 IWEGMICALSYDTSLNNVNPETHKKLTGEGNNNF-EKVIF-GSDSTSLSKSEKPOFLRW 624
QY 226 VTEMGEDPVKRNKKEVLSLKKKCDSCITLRNNGTSNKTCDNENCGACKTQCEKRYKKMER 285
Db 625 LTEWGENFCKEÖKREKYLAKCKDODVDGDG---KC--NGKCVACKDCKÖYHSWIGI 678
QY 286 WKHYSSQKKFKPOLYKNSATYNNGLAVKANSEYTKNDPEYFANSKHAADYLTQLEN 345
Db 679 WIDNKKÖKGRYTEVK-----KIPLYEKDVKNSDD--ARDYLTQLEN 721
QY 346 MICTNG 351
Db 722 MKCVNG 727

```

RESULT 3

US-09-210-288-8
Sequence 8, Application US/09210288
Patent No. 6392026

GENERAL INFORMATION:
APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US

ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288

FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

CITY: Newport Beach

us-10-087-013-9.rai

Page 4

1 NUMBER OF SEQUENCES: 37
2
3 CORRESPONDENCE ADDRESS:
4
5 ADDRESSEE: Knobbe Martens Olson & Bear
6
7 STREET: 620 Newport Center Drive 16th Floor
8
9 CITY: Newport Beach
10
11 STATE: California
12
13 COUNTRY: US
14
15 ZIP: 92660
16
17 COMPUTER READABLE FORM:
18
19 MEDIUM TYPE: Floppy disk
20
21 COMPUTER: IBM PC compatible
22
23 OPERATING SYSTEM: PC-DOS/MS-DOS
24
25 SOFTWARE: Patent In Release #1.0, Version. #1.25
26
27 CURRENT APPLICATION DATA:
28
29 APPLICATION NUMBER: US/09/210,288
30
31 FILING DATE:
32
33 CLASSIFICATION:
34
35 ATTORNEY/AGENT INFORMATION:
36
37 NAME: Fuller, Michael
38
39 REGISTRATION NUMBER: 36,516
40
41 REFERENCE/DOCKET NUMBER: NIH121.1RWDA1
42
43 TELECOMMUNICATION INFORMATION:
44
45 TELEPHONE: (619) 235-8550
46
47 TELEFAX: (619) 235-0176
48
49 INFORMATION FOR SEQ ID NO: 12:
50
51 SEQUENCE CHARACTERISTICS:
52
53 LENGTH: 2710 amino acids
54
55 TYPE: amino acid
56
57 STRANDEDNESS: single
58
59

```

? APPLICANT: Peterson, David S.
? APPLICANT: Su, Xin-zhaun
? APPLICANT: Wellens, Thomas E.
? TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
? NUMBER OF SEQUENCES: 45
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Knobbe Martens Olson & Bear
? STREET: 620 Newport Center Drive 16th Floor
? CITY: Newport Beach
? STATE: California
? COUNTRY: US
? ZIP: 92660
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/487,826B
? FILING DATE: 10-SEP-1993
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Israelson, Ned
? REGISTRATION NUMBER: 29,655
? REFERENCE/DOCKET NUMBER: NIH121.001CPI
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619) 235-8550
? TELEFAX: (619) 235-0176
? INFORMATION FOR SEQ ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 3060 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-487-826B-14

Query Match 32.7%; Score 625.5; DB 2; Length 3060;
Best Local Similarity 38.8%; Pred. No. 9.7e-48;
Matches 137; Conservative 61; Mismatches 126; Indels 29; Gaps 9;

OY 2 CKNVOKLISEQIEKNNIHNCKTDED-AKMKCENTLGEDEGVCMPPRRQNLCHVYL--T 57
DB 1347 CKIVEKILBEKNGRTYVGECPNPKESYPMDCNNIDISHGACMPRRQKLCIYIAHES 1406
58 KLANDSKEDLEAFIKSAAEFPFLRLQYNSKNVEDDKTLHRDMIPPEFFRSMEYTFEGD 117
1407 QTENIKRTDMLKDAFIKTAAEFTFLSWQYKSKNDSEAKILDRGLIPSOFLRSMTYFGD 1466
OY 118 YRDICLDTISEKTIADHDVTTAKKTIYAVQKIGSKTNGKTVLERGMMKEVGLSTWKG 177
DB 1467 YRICLINTDLSKR--QNDVAKAKDKIGKFSKDGSSPSG---LSROEMKKTGPELWKG 1521
OY 178 MLCALS-YNTETKMKDEGVFTYLMKYYIKNNDIKEYLEEFASRPFLRWVTEWGEDFVK 236
DB 1522 MLCALRYVDTDNKRIRKNDYSYDKVNGSQNGSPSLFEPAKPOFLRWMIEMGEERCAE 1581
OY 237 RKKEVLSLKKKCSCLRNNGTSNKTCDNENGCACGTCQEKYKKMMERKKHHYSKKK 236
DB 1582 ROKKENIKDAC-----NEINSTQOCNDAKH--RCNQACRAVOYEENKKEESGQTN 1633
OY 297 FOYKXSATYNNGLAVSEANSFYKNDPEVTEANSKHAHDYLTQLENNICT 349
DB 1634 FVLKANVQPOD-----PEYKGYEKDGVOPIQGN-----EYLLQKCDNNKCS 1675

```

```

? APPLICANT: Chitnis, Chetan
? APPLICANT: Miller, Louis H.
? APPLICANT: Peterson, David S.
? APPLICANT: Su, Xin-zhaun
? APPLICANT: Wellens, Thomas E.
? TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
? NUMBER OF SEQUENCES: 45
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Knobbe Martens Olson & Bear
? STREET: 620 Newport Center Drive 16th Floor
? CITY: Newport Beach
? STATE: California
? COUNTRY: US
? ZIP: 92660
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/487,826B
? FILING DATE: 10-SEP-1993
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Israelson, Ned
? REGISTRATION NUMBER: 29,655
? REFERENCE/DOCKET NUMBER: NIH121.001CPI
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619) 235-8550
? TELEFAX: (619) 235-0176
? INFORMATION FOR SEQ ID NO: 16:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2182 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? FRAGMENT TYPE: N-terminal
? US-08-487-826B-16

Query Match 19.6%; Score 374.5; DB 2; Length 2182;
Best Local Similarity 25.4%; Pred. No. 4.4e-25;
Matches 121; Conservative 54; Mismatches 137; Indels 165; Gaps 19;

OY 1 PCKMVOKLISEQIEKNNIHNCK-----KRTEDAKMKC-----ENTKLGEDGVCMPPR 47
DB 851 PCQIVGKLFPED--DKSLKEACGLKYGGRKEKPPNMKCVTPSGVSTATSGDGAICVPR 908
OY 48 RQNLCHVYL-----TKLANDSKEDLEAFIKSAAEFPFLRLQYNS 89
DB 909 RRLRYVGLISQWASRGDETEVSSSEATSPASQSESKLTALLESAAITFLPMHRYKE 968
OY 90 KNV-----ED--DKTLHRDMIPPEFFRSMEYTFEGDYRDICL--D 124
DB 969 EKKRPATQDAGAGLSLPEPSPGEGDQYLOQTGVLPPLFLRMFTTLADYDILYSGS 1028
OY 125 TDISEKTIADHDVTTAK-----KRTIYAVQKIGSKTTN 156
DB 1029 NDTSD-----TTGKQTPSSNDNLKNIIVLEASGSTEQEKMKQIQAKIKKLNGATS 1081
OY 157 GKAVLER-----EGWKKREYGLSIWGMICALSY-----NTERKK--MDE 193
DB 1082 GVPPIYKNSVKTPOQTWNENIANDIWMAMCALTYSKENDARGTSAKTEQKDKLKAALDE 1141
OY 194 GVR-TYLMKYY-----KNNDIKE--YLEEFASRPFLRWVTEWGEDFVKNRKK 239
DB 1142 ANKNTPLEKYOYTNVNLKLEDSGAKSNDTIQPTPLAKNVELPTFRWLHBMGNSGCFRRK 1201
OY 240 ELVSLKKKCDSCSLRNNGTSNKTCDNENCK-----GACKTQCEKYYKKWME 284

```

```

RESULT 8
US-08-487-826B-16
; Sequence 16, Application US/08487826B
; Patent No. 5993827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.

```

Db 1202 RLALRHEC-----MDEDEKQYSGDEYCEIFSKQYNVLQDLSSSCAKPCLYKWTIE 1256

QY 285 RPKKHYSQKKKKFOLYKNSATYNNGLAVKANSEYTKNDPEYTEANSKADYDKT 341

Db 1257 KKTEYKQKQKAVEQK-----SNTEHQKDKCOTQSN--NAMEFSKT 1298

RESULT 9

US-08-568-459A-6

Sequence 6, Application US/08568459A

Patent No. 5849306

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/568,459A

FILING DATE: 07-DEC-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH121.001CPI

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 749 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium falciparum

US-08-568-459A-6

Query Match 18.5%; Score 353; DB 2; Length 749;

Best Local Similarity 28.2%; Pred. No. 8.9e-24;

Matches 103; Conservative 50; Mismatches 118; Indels 94; Gaps 16;

QY 10 SEQIEKNNNHNC---KTEDAKMKC-ENTKLGDEGVCMPRRQNCVHYLTKLNDKSKE 65

Db 103 SKHEHSSVFGCKTKISKVKKKMKCYSNKNVTKPEGVCGPBRROOLCLGFIPLIR-DGNE 161

QY 66 EDLREAFIKSAAEFTLLRQYNSKNVEDKILHRMIPPEFPRSMYTFGDRDLCIDT 125

Db 162 EGLKDHINKAANEAMHLKEKY--ENAGGDKICN-----AIIGSTADIGDVRGL 209

QY 126 DIEKADHDVYTAKKITAVFQKISGKTNGKRVL-----EREGWKEYGISTWGMCA 181

Db 210 DVM-----RDINT--KLTSEKFORIFMGSGNSRRKKQNDNERRKKMKWKKQNLNLSM-- 259

QY 182 LSVNTEFKKNDSEVRYLIMKIYIKKNDIKELYLEEFASRPFLKRVTEWGDVFYKNNKKEL 241

Db 260 -----VKHLPKGTCKRH--NNEKIPDPLRLKLGWDEFCCEMGTEV 300

QY 242 VSLAKKQSCCLRNNGTSNKTCDDNENGACCKTQCEKTKMKMERKKHYSQKKF----- 297

Db 301 KOLEKICE-----NKNCSEKK-----CKNACSSYEKKIKERKNYINQSKKFPDSK 346

QY 298 QLYKNSATYNNGLAVKANSEYTK-----NDPEYTEANSKADYDKTQLEN 345

Db 347 KLNNKNNLNN-----KFEDSKAYLRSESKQCSNIEFNDETFEPNKKYKA-----C 392

QY 346 MICTN 350

Db 393 MWCEN 397

RESULT 10

US-08-487-826B-6

Sequence 6, Application US/08487826B

Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,826B

FILING DATE: 10-SEP-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH121.001CPI

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 749 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium falciparum

US-08-487-826B-6

Query Match 18.5%; Score 353; DB 2; Length 749;

Best Local Similarity 28.2%; Pred. No. 8.9e-24;

Matches 103; Conservative 50; Mismatches 118; Indels 94; Gaps 16;

QY 10 SEQIEKNNNHNC---KTEDAKMKC-ENTKLGDEGVCMPRRQNCVHYLTKLNDKSKE 65

Db 103 SKHEHSSVFGCKTKISKVKKKMKCYSNKNVTKPEGVCGPBRROOLCLGFIPLIR-DGNE 161

QY 66 EDLEAFIKSAAETFLIRÖYVNSKNVEDDKILHRDIMPPEFFSRMYTFGDYDIDLT 125
Db 162 EGLDHIINKAANYAMHLKERY--ENAGDRIKCN-----ALLGSYADIDIDYRGL 209
QY 126 DISKIDHDVYTKKRTAVFQKIGSKTNGKRVL-----EREGMKREYGLSTIKGMICA 181
Db 210 DVM-----RDINT--NKLSEKFOKIFMGGSRRKKÖNDNERKNMWEKORLWSSM--- 259
QY 182 LSYNTEFKKMDGCVRYTLMKIYYKNNNDIKLEYLEEFASRPPLRWVTWEGEDFYVNRKKEL 241
Db 260 -----VKHPIPKGKTCRHR--NNEFKIPOFLRWLKEWGEFCEMGTEV 300
QY 242 VSLKKKDCSTLRNNGTSNKTCDNENGCACKTQCEKYKKMWERKKHYSSÖKKKF--- 297
Db 301 KÖLEKICE-----NKKCSSEK---CKNACSSYEKWKERKNEYNLÖSKKFFSDK 346
QY 298 ÖLYKNSATYNNGLAVKEANSEYK-----NDPEVTEANSKAKHARDYLUKTÖLEN 345
Db 347 KLNNKNNLYN-----KPEDSKAYLRSESÖKSÖNIEFNDETFPPNKYKEA-----C 392
346 MICTN 350
Db 393 MWCEN 397

RESULT 11
US-09-210-288-6
Sequence 6, Application US/09210288
Patent No. 6392026
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FWDY1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-0176
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 749 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-09-210-288-6

Query Match 18.5%; Score 353; DB 4; Length 749;
Best Local Similarity 28.2%; Pred. No. 8.9e-24;
Matches 103; Conservative 50; Mismatches 118; Indels 94; Gaps 16;
QY 10 SEQLEKNNHNC---KTEDAKWKC-ENTKLGEDEVCMPPRONLCVHYLUKINDSKE 65
Db 103 SKHEHESSVFQCKTKISKVKKKMCCYNNKVTKEPGVGPBRROOLGTYFFLIR-DGNE 161
QY 66 EDLEAFIKSAAETFLIRÖYVNSKNVEDDKILHRDIMPPEFFSRMYTFGDYDIDLT 125
Db 162 EGLDHIINKAANYAMHLKERY--ENAGDRIKCN-----ALLGSYADIDIDYRGL 209
QY 126 DISKIDHDVYTKKRTAVFQKIGSKTNGKRVL-----EREGMKREYGLSTIKGMICA 181
Db 210 DVM-----RDINT--NKLSEKFOKIFMGGSRRKKÖNDNERKNMWEKORLWSSM--- 259
QY 182 LSYNTEFKKMDGCVRYTLMKIYYKNNNDIKLEYLEEFASRPPLRWVTWEGEDFYVNRKKEL 241
Db 260 -----VKHPIPKGKTCRHR--NNEFKIPOFLRWLKEWGEFCEMGTEV 300
QY 242 VSLKKKDCSTLRNNGTSNKTCDNENGCACKTQCEKYKKMWERKKHYSSÖKKKF--- 297
Db 301 KÖLEKICE-----NKKCSSEK---CKNACSSYEKWKERKNEYNLÖSKKFFSDK 346
QY 298 ÖLYKNSATYNNGLAVKEANSEYK-----NDPEVTEANSKAKHARDYLUKTÖLEN 345
Db 347 KLNNKNNLYN-----KPEDSKAYLRSESÖKSÖNIEFNDETFPPNKYKEA-----C 392
346 MICTN 350
Db 393 MWCEN 397

RESULT 12
US-08-568-459A-4
Sequence 4, Application US/08568459A
Patent No. 5849306
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568,459A
FILING DATE: 07-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-0176
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

LENGTH: 1435 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium falciparum
 US-08-568-459A-4

Query Match 17.2%; Score 329; DB 2; Length 1435;
 Best Local Similarity 24.8%; Pred. No. 3.4e-21;
 Matches 87; Conservative 63; Mismatches 113; Indels 88; Gaps 13;

QY 14 EKNNIHCKKTEDAKWKCENKTLGDEGVCMPPRRONLCVHYLTKLNDSS---KEEDLR 69
 DB 463 DKNSV---DTNKTWCKKNPYILSTKDYCVPPRRQELCLGNIDRIYKNLMIKEHILA 518
 QY 70 EAFIKSAAEFTLLRQYINSKKNVEDKILHRDMIPPEFRSMFTYFGDYRDLCLDTDISE 129
 DB 519 IAIYES---RLIKRRYKKNKDDKEVCKIINK-----TFADIRDIIGGTIDYWN 561
 DB 130 KIADHDVTTAKKITAFAVOKIGSKTTNGKKV-----LEREGWKEEGLSIWKGMLCA 181
 DB 562 DLSNRKL-----VGKINTNSKYVHRNKNKDLPRDEWKKYIKKDYWN----- 603
 QY 182 LSYNETKMKDGVATYLMKTIYKKNNDI-KEYLEEFASRPPPLRWVTGEGDFVKNRKE 240
 DB 604 -----VISWFKDKTYCKE--DDIENIPQFFRWFSEMGDDYCODKTKM 644
 QY 241 LVSLKKKDCSCTLRNNGTSNKTCDNENCGACKCTOCEYKKMKMEKKHYSOKKFFOLY 300
 DB 645 IETLAVEC-----KEKCEDDN-----CKSKCNSTKEMISKKEEYKOKAYOYEY 690
 QY 301 KNSATYNNGLAVKEANSEY--KNDPEVTEANSKHAHARDYLTQLENNICT 349
 DB 691 OKGNMYKMYSEKSIKPEVYLKKYSEKCSNLFNFEDEFKEELHSDYKNK-CT 740

RESULT 13

US-08-487-826B-4
 ; Sequence 4, Application US/08487826B
 ; Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
 APPLICANT: Chitnis, Chetan
 APPLICANT: Miller, Louis H.
 APPLICANT: Peterson, David S.
 APPLICANT: Su, Xin-zhaun
 APPLICANT: Wellens, Thomas E.
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/487, 826B
 FILING DATE: 10-Sep-1993

CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Israelson, Ned
 REGISTRATION NUMBER: 29,655
 REFERENCE/DOCKET NUMBER: NH121.001CPI

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1435 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Plasmodium falciparum
 US-08-487-826B-4

Query Match 17.2%; Score 329; DB 2; Length 1435;
 Best Local Similarity 24.8%; Pred. No. 3.4e-21;
 Matches 87; Conservative 63; Mismatches 113; Indels 88; Gaps 13;

QY 14 EKNNIHCKKTEDAKWKCENKTLGDEGVCMPPRRONLCVHYLTKLNDSS---KEEDLR 69
 DB 463 DKNSV---DTNKTWCKKNPYILSTKDYCVPPRRQELCLGNIDRIYKNLMIKEHILA 518
 QY 70 EAFIKSAAEFTLLRQYINSKKNVEDKILHRDMIPPEFRSMFTYFGDYRDLCLDTDISE 129
 DB 519 IAIYES---RLIKRRYKKNKDDKEVCKIINK-----TFADIRDIIGGTIDYWN 561
 DB 130 KIADHDVTTAKKITAFAVOKIGSKTTNGKKV-----LEREGWKEEGLSIWKGMLCA 181
 DB 562 DLSNRKL-----VGKINTNSKYVHRNKNKDLPRDEWKKYIKKDYWN----- 603
 QY 182 LSYNETKMKDGVATYLMKTIYKKNNDI-KEYLEEFASRPPPLRWVTGEGDFVKNRKE 240
 DB 604 -----VISWFKDKTYCKE--DDIENIPQFFRWFSEMGDDYCODKTKM 644
 QY 241 LVSLKKKDCSCTLRNNGTSNKTCDNENCGACKCTOCEYKKMKMEKKHYSOKKFFOLY 300
 DB 645 IETLAVEC-----KEKCEDDN-----CKSKCNSTKEMISKKEEYKOKAYOYEY 690
 QY 301 KNSATYNNGLAVKEANSEY--KNDPEVTEANSKHAHARDYLTQLENNICT 349
 DB 691 OKGNMYKMYSEKSIKPEVYLKKYSEKCSNLFNFEDEFKEELHSDYKNK-CT 740

RESULT 14

US-09-210-288-4
 ; Sequence 4, Application US/09210288
 ; Patent No. 6392026

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
 APPLICANT: Chitnis, Chetan
 APPLICANT: Miller, Louis H.
 APPLICANT: Peterson, David S.
 APPLICANT: Su, Xin-zhaun
 APPLICANT: Wellens, Thomas E.
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US
 ZIP: 92660

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/210, 288
 FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-09-210-288-4

Query Match 17.2%; Score 329; DB 4; Length 1435;
Best Local Similarity 24.8%; Pred. No. 3.4e-21;
Matches 87; Conservative 63; Mismatches 113; Indels 88; Gaps 13;

QY 14 EKNNHNCCKTDEDAKCKCENTKLGEDGVCMPPRONLCVHYLTKLNDSD---KEEDLR 69
DB 463 DKNSV---DTMTKWECKNPYILSTKDYCPPRRELCLGNDRYDKNLMLKEHILA 518
QY 70 EAFIKSAAEFTLLROYNSKNVEDKILHRMIPPEFRSMYTGVDRLDCLDIDISE 129
DB 519 IAIYES---RIUKRIRYKDKDEVCITNK-----TFADIRDIIGGTDVYN 561
QY 130 KIADHDVTTAKKKTITAVFOKIGSKTNGKRV-----LEREGMKKEGLSTIMKMLCA 181
DB 562 DLSNRRL-----VGKINTNSKYVHRNKKNDLFRDEMKVYIKKQVWN----- 603
QY 182 LSTNITETKMDGCVTYLTKYIKKNDI-KEYIEFPASRPPLRWVTEGDEDFVKNRKE 240
DB 604 -----VISWFKDKTYCKE--DLEINIPQEFRRFSEMGDDYCDQKTKM 644
QY 241 LVSLKKKCDSCITLRNNGTSNKTCDNENGACCKTOCEKRYKKMERKKHYSQKKKFOLY 300
DB 645 IETLAKYEC-----KEKPEDDN---CKSKCNSTKEMVSKKEEKNQAKOYOEY 690
QY 301 KNSATYNGGLAVKEANSEYTY--KNDEPVTANSASAKHARDYLTQLENMICT 349
DB 691 QKGNNTKMYSEFKSIKPEYVLLKYSSEKCNLNFEDFEKLHSDYKMK-CT 740

RESULT 15
US-08-568-459A-17
Sequence 17, Application US/08568459A
Patent No. 5849306
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chltnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/568,459A
FILING DATE: 07-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 324 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-568-459A-17

Query Match 14.8%; Score 284; DB 2; Length 324;
Best Local Similarity 29.4%; Pred. No. 5.3e-18;
Matches 74; Conservative 15; Mismatches 145; Indels 18; Gaps 6;

QY 43 CMPPRONLCVHYL--TKLNDSEKEDLRFAFIKSAAEFTLLROYNSKNVEDKILHR 100
DB 25 CIPPRROKLCILHYLXXX 84
QY 101 DMIPPEFRSMETPEDYRDICLDPDISKILADHDVTTAKKKTITAVFOKIGSTTGKRV 160
DB 85 XXXXXDFKQMFYTFADYRDICLGTDISK---KDTISXXXXXXXXXXXXXXXXISN--SI 139
QY 161 LEREGMKKEGLSTIMKMLCAL-----SYNFETKMDGCVRTYLMKYIKKNDIKEYLE 214
DB 140 RYKSKMWTNGPIYTWGMLCALXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 199
QY 215 EFASRPPLRWVTEGDEDFVKNRKKELVSLKKKCDSCITLRNNGTSNKTCDNENGACAKT 274
DB 200 XXXXRPQFLRWLTWENGECCKEKKYVLLAK---CXXXXXXXXC--XXXXVACND 254
QY 275 QCEKRYKKMERW 286
DB 255 QCKYHWSWIGIW 266

Search completed: June 20, 2003, 15:07:02
Job time: 20.1745 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 15:05:30 ; Search time 36.7511 seconds
(without alignments)
1033.453 Million cell updates/sec

Title: US-10-087-013-9

Perfect score: 1913
Sequence: 1 PCKWQKLISEQIEKNININ.....AKHARDYLTQLENNICTNG 351

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Published_Applications_AA.*
2: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	646	33.8	921	US-10-153-273-8	Sequence 8, Appl
2	625.5	32.7	2710	US-10-153-273-12	Sequence 12, Appl
3	353	18.5	749	US-10-153-273-6	Sequence 6, Appl
4	353	18.5	1086	US-09-924-154-15	Sequence 15, Appl
5	333.5	17.4	1143	US-09-924-154-14	Sequence 14, Appl
6	329	17.2	1135	US-10-153-273-4	Sequence 4, Appl
7	312	16.3	1421	US-09-924-154-13	Sequence 13, Appl
8	284	14.8	324	US-10-153-273-17	Sequence 17, Appl
9	272.5	14.2	311	US-10-153-273-21	Sequence 21, Appl
10	269	14.1	1115	US-10-153-273-2	Sequence 2, Appl
11	231	12.1	700	US-10-153-273-10	Sequence 10, Appl
12	226.5	11.8	972	US-09-924-154-16	Sequence 16, Appl
13	220	11.5	1501	US-09-924-154-17	Sequence 17, Appl
14	171	8.9	277	US-10-153-273-15	Sequence 15, Appl
15	144.5	7.6	411	US-10-153-273-20	Sequence 20, Appl
16	137	7.2	282	US-10-153-273-16	Sequence 16, Appl
17	135	7.1	291	US-10-153-273-13	Sequence 13, Appl
18	131.5	6.9	362	US-10-153-273-18	Sequence 18, Appl
19	101	5.3	1167	US-09-815-242-11522	Sequence 11522, A

20	101	5.3	1938	9	US-10-171-311-164	Sequence 164, App
21	101	5.3	1945	9	US-09-927-597-2	Sequence 2, Appl
22	1972	5.3	1972	9	US-10-171-311-162	Sequence 162, Appl
23	101	5.3	1979	9	US-09-927-597-4	Sequence 4, Appl
24	99	5.2	874	9	US-09-893-519A-50	Sequence 50, Appl
25	98.5	5.1	1092	9	US-09-423-126-5	Sequence 5, Appl
26	98	5.1	380	10	US-09-134-333-13	Sequence 13, Appl
27	98	5.1	456	10	US-09-925-300-1595	Sequence 1595, Ap
28	98	5.1	1805	9	US-09-820-843A-73	Sequence 73, Appl
29	97.5	5.1	665	9	US-09-820-843A-107	Sequence 107, App
30	97	5.1	208	9	US-10-084-205-52	Sequence 52, Appl
31	97	5.1	208	10	US-09-925-637-52	Sequence 52, Appl
32	97	5.1	411	9	US-10-153-273-19	Sequence 19, Appl
33	97	5.1	711	9	US-09-298-523B-3	Sequence 3, Appl
34	95	5.0	677	10	US-09-745-763-168	Sequence 168, App
35	95	5.0	716	10	US-09-815-242-12403	Sequence 12403, A
36	95	5.0	1475	10	US-09-740-274-2	Sequence 2, Appl
37	94.5	4.9	271	9	US-10-153-273-14	Sequence 14, Appl
38	94.5	4.9	2353	10	US-09-797-862-33	Sequence 33, Appl
39	93	4.9	380	10	US-09-134-333-12	Sequence 12, Appl
40	93	4.9	654	10	US-09-952-013A-5	Sequence 5, Appl
41	91	4.8	315	9	US-10-106-698-5557	Sequence 5557, Ap
42	90	4.7	610	9	US-09-802-640-36	Sequence 36, Appl
43	90	4.7	621	9	US-10-205-823-357	Sequence 357, App
44	89.5	4.7	621	10	US-09-856-247A-2	Sequence 2, Appl
45	89.5	4.7	670	9	US-09-298-523B-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1
US-10-153-273-8
Sequence 8, Application US/10153273
Patent No. US20020169305A1
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
Chitnals, Chetan
Miller, Louis H.
Peterson, David S.
Su, Xin-zhaun
Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/153,273
FILING DATE: 21-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:

LENGTH: 921 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
SEQUENCE DESCRIPTION: SEQ ID NO: 8
US-10-153-273-8

Query Match 33.8%; Score 646; DB 9; Length 921;
Best Local Similarity 41.0%; Pred. No. 5,3e-45;
Matches 150; Conservative 54; Mismatches 112; Indels 50; Gaps 16;

QY 2 CKMYOKLISEQIEKNNHNCCKTED-AKKCENT--KIGDEGVCMPPRONLCVHYLT 58
DB 396 CEIYAEMLKDNKNGRTYGECKRYKTESEMTCDSEKIKNGQ-HGACIPPRRCKLHYLEK 454
QY 59 LNDKSEEDLEAREFKSAAEFTLLROY--NSKNVED-DKILHRDMLPEPFRSMET 114
455 IMTNTNE--LKAFKCAAAETFLMOMYKKDKNGMNEDELDEKLGKGIPEDFRQMFYT 512
QY 115 FGDRDLCIDPDI-SEKIAHDVTTAKKKTAVQKIGSKTTNGKVKYLEREGWREXGLS 173
DB 513 FADYRDICLGDISSKDKTSKGVKCNIDVFKYISN-----SIRYKSMWETNGPV 566
QY 174 IMKMGALASYNT-----ET-KKMDGVRTYLMKYIKNDIKETYLEEASRPPLRW 225
DB 567 IMEMLCALSYDITSLNWNVPETHKRLKLEGNNE-EKVIYF-GSDSSTLLSFSEPOPLRW 624
QY 226 VTEMGEDPVKRNKKELVSLKKCDSCITLRNNGTSNKTCDNENCGACKTOCEKRYKKMER 285
DB 625 LTEMGEFNECKQEKKEYKVLAKCKDVGDDG---KC--NGKCVACKDCKOYHSHIGI 678
QY 286 MKKYSOKKKFOLYKNSATYNNGLAVKEANSEYTKNDPEVTEANSKHAHDYLTQLEN 345
DB 679 WIDNYKKOKRGYEVK-----KIPLYKEDKDVKNSSD---ARDYLKTQLEN 721
QY 346 MICTNG 351
DB 722 MKCVNG 727

RESULT 2
US-10-153-273-12
Sequence 12, Application US/10153273
Patent No. US20020169305A1
GENERAL INFORMATION:
APPLICANT: Slim, Kim L.
Chitnis, Chetan
Miller, Louis H.
Peterson, David S.
Su, Xin-zhaun
Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/153,273
FILING DATE: 21-May-2002

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
SEQUENCE DESCRIPTION: SEQ ID NO: 12
US-10-153-273-12

Query Match 32.7%; Score 625.5; DB 9; Length 2710;
Best Local Similarity 38.8%; Pred. No. 1.1e-42;
Matches 137; Conservative 61; Mismatches 126; Indels 29; Gaps 9;

QY 2 CKMYOKLISEQIEKNNHNCCKTED-AKKCENTKIGDEGVCMPPRONLCVHYLT--T 57
DB 1349 CKIYERILSEKNGRTYGEENPKESYTPMDCKNNIDISHDGACMPRRCKLIYIAHSS 1408
QY 58 KLNDKSEEDLEAREFKSAAEFTLLROYYNNSKNVEDDKILHRDMLPEPFRSMETFGD 117
DB 1409 QTEINIKTDNLKAPFTAAETFLSMQYKSKNDSEAKILDRGLIPSOFLSMYTFED 1468
QY 118 YRDICLDTISEKIAHDVTTAKKKTITAVFOKIGSKTTNGKVKYLEREGWREXGLS 177
DB 1469 YRDICLNTDISK--ONDAKAKDKTGFEFSKDGSKSPSG---LSRQEWKKTNGEIMWG 1523
QY 178 MCAIS--YNTETKKMDGVRTYLMKYIKNDIKETYLEEASRPPLRWTEMGEDPVKN 236
DB 1524 MCALFKYVTDIDNKKRIKNDYSYDKVNSQNGNPSLEEFKAPKPOFLRMIMGEFPCAE 1583
QY 237 RKKEVLSLKKCDSCITLRNNGTSNKTCDNENCGACKTOCEKRYKKMERKHYSSQKK 296
DB 1584 ROKKEKITIDAC-----NEINSTOCCNDAKH--RCNACRAYOEVENKKKEFGQTNN 1635
QY 297 FOLYKNSATYNNGLAVKEANSEYTKNDPEVTEANSKHAHDYLTQLENMIGT 349
DB 1636 FYLKANVQPD-----PEYKGYEYKDGVOPIQGN-----EYLLOKCDNNKCS 1677

RESULT 3
US-10-153-273-6
Sequence 6, Application US/10153273
Patent No. US20020169305A1
GENERAL INFORMATION:
APPLICANT: Slim, Kim L.
Chitnis, Chetan
Miller, Louis H.
Peterson, David S.
Su, Xin-zhaun
Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US

```

? TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
? FILE REFERENCE: 05213-0465 43170-262105
? CURRENT APPLICATION NUMBER: US/09/924,154
? PRIOR FILING DATE: 2001-08-07
? PRIOR APPLICATION NUMBER: US 60/223,525
? PRIOR FILING DATE: 2000-08-07
? NUMBER OF SEQ ID NOS: 17
? SOFTWARE: Patentin version 3.1
? SEQ ID NO 15
? LENGTH: 1086
? TYPE: PRT
? ORGANISM: Mammalian
US-09-924-154-15

Query Match          18.5%; Score 353; DB 10; Length 1086;
Best Local Similarity 28.2%; Pred. No. 1,1e-20;
Matches 103; Conservative 50; Mismatches 18; Indels 94; Gaps 16;

OY      10 SEQIEKNINICK---KTEDAKWC-ENTLIGEDGYCMPRRONLCVHLYTKINDSKE 65
        |:|:::| | | | | | | | | | | | | | | | | | | | | | | | | |
Db       345 SKEHEESVGCCTKTISKVKKKWNCYNNNKYTPKEGCGPPRROOLCLGYIFILIR-DGNE 403
        ||:::| | | | | | | | | | | | | | | | | | | | | | | | | |

OY      66 EDLRERATIKSAAEFTLLROYNSKNKNEDDKILLRDIMPEEPFRSMYTFGDYDIDLDT 125
        |:|::| | | | | | | | | | | | | | | | | | | | | | | | | |
Db       404 EGLDDHINKANVEAMHLKERY--ENAGGOKICN-----ALLGSYADIGDIVRGL 451
        |:|::| | | | | | | | | | | | | | | | | | | | | | | | | |

OY      126 DISRKIDHDVTAAKKKITAVFOKIGSKTTNGKKVL----EREGRMKREYSIWKGMICA 181
        |:|::| | | | | | | | | | | | | | | | | | | | | | | | | |
Db       452 DVM-----RINT--NKLSRFQITPGWGCSNRKRKODNNERNKRWMEKQRULLISSM--- 501
        |:|::| | | | | | | | | | | | | | | | | | | | | | | | | |

OY      182 LSYNETKKMGDEGVRIYLMKIYKNNDIKEYLEEFASRPPLRVATTEMGEDFYVNRRKEL 241
        |:|::| | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

OY      242 VSLKKKCDSCITLRRNGTSNKTCCDNEGCAGCTOCEYKKWMEKWKHYSSQKKF----- 297
           ||| :      ||| :      ||| :      ||| :      ||| :      ||| :
Db      543 KOLEKICE-----NNKCSSEKK-----CKNACSSYEYKWKERKKNEMYLSQSKKFPDSK 588
OY      298 OIVKNSATYNNGLAVKEANSEYTK-----NDPEYTEKNSAKHARDYLKTOLEN 345
           ||| :      ||| :      ||| :      ||| :      ||| :      ||| :
Db      589 KLKKNNKLYN-----KEDSKAYLRSESKOCSNIEFNDEFTFPNNKYKEA-----C 634
OY      346 MICTN 350
           ||| :      ||| :
Db      635 MVCEN 639

RESULT 5
US-09-924-154-14
; Sequence 14, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
; APPLICANT: Natum, David L.
; APPLICANT: Slim, Kim L.
; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
; FILE REFERENCE: 05213-0465 43170-262105
; CURRENT APPLICATION NUMBER: US/09/924,154
; PRIOR FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/222,525
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Version 3.1
; SEQ ID NO 14
; LENGTH: 1143
; TYPE: PRT
; ORGANISM: Mammalian
US-09-924-154-14

Query Match 17.4%; Score 333.5; DB 10; length 1143;
Best Local Similarity 27.9%; Pred. No. 4,8e-19;
Matches 92; Conservative 46; Mismatches 131; Indels 61; Gaps 11;

Y      9 ISDOIENNNH---NCKKTEDAKMKCENTKIGEDGVCMPPRONLCVHYITKILNDSK 64

```

```

Db      448 IASQINVNDLRGCGCNKSNNEKSMNCTGFTFNKFPCTCEPPRQTLCLG-RTYLLHRGH 506
        65 EEDLREAFIKSAAEFTLLROYNSKNVEDDKILHRMIPPEFRSRFEYEGDROICLD 124
        507 EEDYKELLASIEYEAOLKRYKER--DENAL-----CSTIONSYADLADIKG 554
        125 TDISEKADHDVTTAKKITAVFOKIGSKTTNGKYLE--REGWMEYGLSTMGMLCAL 182
        555 SDI---IKDY-----YGKMEENLNKYNKDKKRNESLKITREKWMMDKENVMK----- 601
        183 SYNTETKKMDEGVRTYLMKTIYKNNDIKEYLEEFASRPPLRWTEGDEFFVKNRKKELY 242
        602 -----VMSAVLNKKECTCKDYDFQKIPQFLRWFKEGDDDFCEKREKRIY 645
        243 SLKKKDCSTLRNNGTSNKTCDNENCGACKTQCEYKKMMERWKHHYSOKKKFOLYKN 302
        646 SESEFKVEC-----KKKCCDEN---TKNKCSEYKKWMDLKSEYKQVODKYTKDN 694
        303 SATYNN--GLAVEANSEYTKNDPEVTEAN 330
        695 KKMVDNIDEVKNKEANYLKEKSECKDVN 724

```

RESULT 6

```

US-10-153-273-4
Sequence 4, Application US/10153273
Patent No. US20020169305A1
GENERAL INFORMATION:

```

APPLICANT: Sim, Kim L.

Chilens, Chetan
Miller, Louis H.
Peterson, David S.
Su, Xin-zhaun
Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VITAX

AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobe Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/153,273

FILING DATE: 21-May-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/210,288

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Fuller, Michael

REGISTRATION NUMBER: 36,516

REFERENCE/DOCKET NUMBER: NIH121.1FMDV1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1435 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

```

ORGANISM: Plasmodium falciparum
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-153-273-4

```

```

Query Match      17.2%; Score 329; DB 9; Length 1435;
Best Local Similarity 24.8%; Pred. No. 1,5e-18;
Matches 87; Conservative 63; Mismatches 113; Indels 88; Gaps 13;

```

```

        14 EKNNHCKTTEADAKKCCENTKLGEDGVCMPRRQNLCHYITKLNDSS-----KEEDLR 69
        463 DKNSV-----DTNTRWCKKNPILSTKDVCPVPRROELCIGNIDRIYDKMLMIKEHILA 518
        70 EAFIKSAAEFTLLROYNSKNVEDDKILHRMIPPEFRSRFEYEGDROICLDIDISE 129
        519 IATYES-----RLIKRKKNNDDKEVCITIK-----TFADIDIDIGGIDYWN 561
        130 KADHDVTTAKKITAVFOKIGSKTTNGKVV-----LERGWMKEYGLSTMGMLCA 181
        562 DLSNRKL-----VGRINTSKYVHNRKKNKDLFRDWMVKVIRKDVWN----- 603
        182 LSYNTEKKMDEGVRTYLMKTIYKNNDI--KEYLEEFASRPPLRWTEGDEFFVKNRKE 240
        604 -----VISWVERDKTVCKE--DQIENIPQFRFSEWGDYCCQDKTKM 644
        241 LVSLKKKDCSTLRNNGTSNKTCDNENCGACKTQCEYKKMMERWKHHYSOKKKFOLY 300
        645 IETLKVEC-----KKPCEDDN---CKSKNCSYKEMWISKKKEEYNKAKQOYQY 690
        301 KNSATTYNGLAVEANSEY--KNDEPVTEANSAKHAADYLTQLENMICT 349
        691 QKGNMYKMYSEFKSIKPEVYLKRYSEKSNLNEDEFKEELHSDYKNK--CT 740

```

RESULT 7

```

US-09-924-154-13
Sequence 13, Application US/09924154
Patent No. US20020127241A1
GENERAL INFORMATION:

```

APPLICANT: Narum, David L.

FILE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use

TITLE REFERENCE: 05213-0465 43170-262105

CURRENT APPLICATION NUMBER: US/09/924,154

CURRENT FILING DATE: 2001-08-07

PRIOR APPLICATION NUMBER: US 60/223,525

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 17

SOFTWARE: Patentin version 3.1

SEQ ID NO 13

LENGTH: 1421

TYPE: PRT

ORGANISM: Mammalian

US-09-924-154-13

```

Query Match      16.3%; Score 312; DB 10; Length 1421;
Best Local Similarity 24.2%; Pred. No. 3.8e-17;
Matches 85; Conservative 61; Mismatches 117; Indels 88; Gaps 13;

```

```

        14 EKNNHCKTTEADAKKCCENTKLGEDGVCMPRRQNLCHYITKLNDSS-----KEEDLR 69
        463 DKNSV-----DTNTRWCKKNPILSTKDVCPVPRROELCIGNIDRIYDKMLMIKEHILA 518
        70 EAFIKSAAEFTLLROYNSKNVEDDKILHRMIPPEFRSRFEYEGDROICLDIDISE 129
        519 IATYES-----RLIKRKKNNDDKEVCITIK-----TFADIDIDIGGIDYWN 561
        130 KADHDVTTAKKITAVFOKIGSKTTNGKVV-----LERGWMKEYGLSTMGMLCA 181
        562 DLSNRKL-----VGRINTSKYVHNRKKNKDLFRDWMVKVIRKDVWN----- 603
        182 LSYNTEKKMDEGVRTYLMKTIYKNNDI--KEYLEEFASRPPLRWTEGDEFFVKNRKE 240
        604 -----VISWVERDKTVCKE--DQIENIPQFRFSEWGDYCCQDKTKM 644

```

```

Db      85 XXXXXDFKRMQMTFDFDYLICGDTPISSK---KOTSSXXXXXXXXXXXXXXXXXKISN--SI 139
QY      161 LEREGWKEVGLSINKGMLCAL-----SYNETKKMDECVRTYLMKYYIKNNDIKEYLE 214
      140 RYKSMWENENGPIWEGMLCALXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 199
QY      215 EFASRPPELRMTWENGEDFVKNRKKELVSLKKKDCSTLRNNGTSNKTCDONENGCACKT 274
      200 XXXXRPQPLRMTWENGENFCCKEKKYKVLAK---CXXXXXXXXXXXXXC--XXXCVCACKD 254
Db      275 QCEKYKKMERW 286
QY      255 QCKQYHSHWIGIW 266
Db

RESULT 9
US-10-153-273-21
: Sequence 21, Application US/10153273
: Patent No. US20020169305A1
: GENERAL INFORMATION:
: APPLICANT: Sim, Kim L.
: Chitnis, Chetan
: Miller, Louis H.
: Peterson, David S.
: Su, Xia-zhaun
: Wellem, Thomas E.
: TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINSS
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Knodde Martens Olson & Bear
: STREET: 620 Newport Center Drive 16th Floor
: CITY: Newport Beach
: STATE: California
: COUNTRY: US
: ZIP: 92660
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/153,273
: FILING DATE: 21-May-2002
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/09/210,288
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Fuller, Michael
: REGISTRATION NUMBER: 36,516
: REFERENCE/DOCKET NUMBER: NIH121.1FMDV1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 235-8550
: TELEFAX: (619) 235-0176
: INFORMATION FOR SEQ ID NO: 21:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 311 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: internal
: ORIGINAL SOURCE:
: SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-10-153-273-21

Query Match      14.28; Score 272.5; DB 9; Length 311;
Best Local Similarity 28.48; Pred. No. 1e-14;
Matches 71; Conservative 22; Mismatches 140; Indels 17; Gaps 6;

```


Query Match	28.6%	Score 547.5;	DB 2;	Length 1711;
Best Local Similarity	37.3%;	Pred. No. 6.2e-30;		
Matches 128; Conservative	53;	Mismatches 105;	Indels 57;	Gaps 14;

[illegible]

OY 332 AKH-ARDYLKTOLE 344
 Db 1097 AONSATEEYKNIE 1110

RESULT 5

variant-specific surface protein 3 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T28625

R:Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herfelde, J.A.; Peterson, D.S.

Cell 82, 89-100, 1995

A:Title: The large diverse gene family var encodes proteins involved in cytoadherence at

A:Reference number: Z20487; MID:95330813; PMID:7606788

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Accession: T28625

A:Molecule type: DNA

A:Cross-references: EMBL:LA0609; MID:9886376; PID:9886377; PIDN:AA05397.1

A:Gene: var-3

A:Introns: 2597/3

Query Match 26.2%; Score 501; DB 2; Length 3006;

Best Local Similarity 30.0%; Pred. No. 2e-26;

Matches 125; Conservative 63; Mismatches 123; Indels 106; Gaps 14;

OY 1 PCKMVKL-----ISEQIEKNNIHNCKTEDAKMKCENTKLGEDGVCMPRRONLCVH 54

Db 1755 PDIIVKTLFTTETLEACPTKYVNGREKPPN--WKICISGSDASGICIPRRRLYLH 1812

OY 55 YL-----TKLNDSEK-----EEDLRAPFKSAAEFPL-RQYNSKNVE----- 93

Db 1813 KIEGVDTVSSDDETTPTTHDALREAFIQTAAVEFFELHWRKIKIKERQELONGTF 1872

OY 94 -----DDKILHRDMIPEEFRRSMFTFGDYRDIC-----LDTOL 127

Db 1873 LLEPAOKVSPEDNPEHPQKKLEKGIPEEFKROMFTLIGDYRDLGVKADVAQALEASG 1932

OY 128 SEKIADHDVTTAKKITAVAFQIKGSKTTNGK---VLEREGWKEVGLSIWGMICALS 183

Db 1933 DNKSGDKNIKIDISEKIVSEKSGEQTTPGPKPGQTTPKPEEWOKNGEIHMANICALT 1992

OY 184 YNFKETKMDGVFTYLM-----KITYKNNDIKY-----LEEFASRPF 222

Db 1993 HNTDTRQVDDQVKGOLFENGNKTPKNSQYQYKNTISSVSSNGPIGNIKLEQFASRPTF 2052

OY 223 LRWVTEWGEDFVNRKKELVSLKKKDCSTLRNNGTSNKTCDNE----- 267

Db 2053 LRHLEMGEEFCCKQTHKLEKEEC-----HKDGNRNCDDDDGFECKECKCKDGSFE 2105

OY 268 --NCGACKTCCEKYYKKMERMWKKRHSYSSQKKRFLYKNSATYNNGLAVKEANSETYKN 322

Db 2106 TLKCPSCAKSKCYKWKISRKDEFTKQGAVEKQKADAGNNN-DYKEP-SKTLRN 2160

RESULT 6

erythrocyte membrane protein 1 - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T28157

R:Yang, S.O.; Wollish, W.S.; Gut, J.; Wu, J.; Ahn, J.; Petersen, C.; Fujioaka, H.; Alkawa

submitted to the EMBL Data Library, July 1995

A:Description: The molecular cloning and DNA sequence analysis of Plasmodium falciparum

A:Reference number: Z20479

A:Accession: T28157

C:Genetics:

A:Gene: EMP1

Query Match 23.2%; Score 443; DB 2; Length 2212;

Best Local Similarity 26.2%; Pred. No. 1.5e-22;

Matches 128; Conservative 59; Mismatches 133; Indels 168; Gaps 19;

OY 8 LISEQIEKNNIHNCKTE-----DAKMKC-----ENTKLGED----- 39

Db 1272 IVAGVLTRENTLEACPTKYCPKAPTSMKCIPTKTAANAGSESSGNGALORAKRATVES 1331

OY 40 -----EGVCMPPRRONLCVHLYRLND-----DSKEE 66

Db 1332 GSPVTSNSGICIPRRRLR---YIOKLHDMAGNTVSGAQTPOGTSPPSGKETPSD 1388

OY 67 DLREAFIKSAATFPL-RQYNSKNVE-----DD 95

Db 1389 KLRTAFIQSAIETFFLMDRYKKEKEKEKRVANGGLVPSLNGSPPOGVTGDSPOS 1448

OY 96 KILHRDMIPEEFRRSMFTFGDYRDICL--DPTDISKIDHDVTTAKKITAVAFQKI-- 150

Db 1449 KLOQTGVIPPEFLRQMFYTLGDIADIFFGKNDYIOTKNGDKDIAREKKIKAIERYLK 1508

OY 151 --GSKTTNGKRVLEREGWKEVGLSIWGMICALSYNTETK-----KMDGVFTYL--- 199

Db 1509 NADSPPEDEK---RQWMDQNGEHIWNGICALTYKERDEKGPLKQNGLSALMDEK 1565

OY 200 -----MKIY-----KND--IKELYEFASRPFPLRWVTEWGEDFVKNRK 238

Db 1566 NKRPKDOKYDYDKLDENSGTSPKTNHDVPTPLTFISRPYFYLEWGTGTFPCRERK 1625

OY 239 KELVSLKKKDCSTLRNNGTSNKTCDNEN-----CGACKTCCEKYYKKV 282

Db 1626 KRLKIKYEC---MDEDGKKQKCSGDGECEELRKQDYSTVNDFYCPREGKRCRFKRW 1681

OY 283 MBRWKHHYSQKKRFLYKNSATYNNGLAVKEANSETYKNMDEPTEANSKHAARDYLKTO 342

Db 1682 IEKKIDEDYDKQKRAYNNQKTDARRNN-----NDNAFS-----TTLDTCTTAGDPLQT- 1728

OY 343 LENMICTN 350

Db 1729 LKNGPCKN 1736

RESULT 7

variant-specific surface protein 1 homolog 3D7var1 - malaria parasite (Plasmodium fal

C:Species: Plasmodium falciparum

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T18399

R:Rubio, J.P.; Thompson, J.K.; Cowman, A.F.

EMBO J. 15, 4069-4077, 1996

A:Title: The var genes of Plasmodium falciparum are located in the subtelomeric regio

A:Reference number: Z18927; MID:96324414; PMID:8670911

A:Accession: T18399

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2042 <RUB>

A:Cross-references: EMBL:U05324; MID:91297090; PID:91297091; PIDN:AA09769.1

A:Map position: 12

C:Genetics:

A:Gene: EMP1

Query Match 22.7%; Score 433.5; DB 2; Length 2042;

Best Local Similarity 25.1%; Pred. No. 6.1e-22;

Matches 131; Conservative 68; Mismatches 142; Indels 181; Gaps 19;

OY 1 PCKMVKLISEQIEKNNIHNCKTEDAKMKC-----ENTKLG----- 37

Db 695 PCKIVEELFKSTKFEADACGLKYKNGKVCVHHTSDKGSPTARSHVARSADGAPSG 754

OY 38 -EDGVCMPRRONLCVHLYRLNDSEKEDLREAFIKSAATFPLRQYNS----- 89

Db 755 DKDGAICPPRRKRLYLHIEGV-DTDDKSLRKMFESAAVETFLMDRKYKLTTPQSG 813
QY 90 -----KNED-----DKILHRMIPPEFRSMFYEGYRDICLDTISEKA 132
Db 814 SPLLGWITGVGVENDDENNPEKILQKEIPDGLRQMFYLLGDVRLDLSGDDKKNG 873
QY 133 DHDVYTA-----KKKITAVFQKIGSKTTNGK-----VLERGWMKEYGLSIWKM 178
Db 874 YSDIVSGDNVIRENTIEKTIASFQNGKEGTPHVPKNPYOTQPMKDNKAHIMQM 933
QY 179 LCAISYNET-----KKMDEGVARY-----LKKYIKNND 208
Db 934 ICALVYEETSGSDEDEKTEKDADAYKKFKFPNGNPDPQCTNGTSNEPIISOYQDYV 993
QY 209 IKE-----YLEFASRPFLRWYEMWGEDVFNKKEKELVSLK 245
Db 994 LKEENNAGMSTSPKSTASAPSDNPTTLTQFLRPTLYRLEWMOQNFCEKKEKRLAOY 1053
QY 246 KKCDSCITLRNNTSKTCD-----DNEGCACTQCEKYKMMERMK 288
Db 1054 EDC-----RGN-----DKVCSGDGEDCEVVRKODYSKISNFGCGRCRKYKNWIKRT 1105
QY 289 HYSQKRRFQLYKNS-ATYNNG-----LAVKANSEY-----KN- 322
Db 1106 EFDECKKAVDRKDKYTKENKGAESKHSISQNTVYKMLGTDYASIESFLEKLGSCSKNN 1165
QY 323 -----DPEVT--EANSAKHARDYLTQLENMICTNG 351
Db 1166 KDNGDGTINFKEPDYTFKPADCKPCSEF-KVCENGCOSG 1206

RESULT 8
T18378
variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
N:Alternate names: erythrocyte membrane binding protein 1 (EMP1)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T18378
R:Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bl, X.; Ma, X.C.; Feldman, M.; Taraschl, T.
Cell 82, 77-87, 1995
A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and
A:Reference number: 218925; MUID:95330812; PMID:7541722
A:Accession: T18378
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2924 <R>
A:Cross-references: EMBL:U27338; NID:914918; PID:914919; PIDN:AAB60251.1
C:Genetics:
A:Gene: EMP1
A:Introns: 2476/3

Query Match 22.6%; Score 432; DB 2; Length 2924;
Best Local Similarity 26.3%; Pred. No. 1.2e-21;
Matches 132; Conservative 59; Mismatches 134; Indels 176; Gaps 19;

QY 2 CKMWQKLISEQIKNNHNCCTEDAKKC-----ENKLGDEG-VC 43
Db 839 CSYDKALKGKLDACLTGKTAPTSMKCPISGNNTTSTTKPGAAGPSSGKDTGSIC 898
QY 44 MPRRQMLCVHYLTKND-----LROYNSKN-----VEDDKIL 98
Db 899 VPPRRRL-----YVGLKLDWAGGETTEAKSQETSGGQKTPSGNESSEPKLPQCTPTTK 955
QY 62 DSKEDLEAFIYSAAEFL-----LROYNSKN-----VEDDKIL 98
Db 956 ETPESLILHAFVSPFLRRFLPMHFKFEQWKAQAGATGQOTLIGLDDGGERTPKL 1015
QY 99 HRDMPPEFRSMFYEGYRDICL-DTDL-----SEKIAHDVTA-KKITAVFQKIGS 152
Db 1016 KTHGIPDFLRQMFYLLGDVRLDILVHTSGKEDMOIMEALOKRIEOTLPTSGS 1075
QY 153 -KTNGKYLE--REGWMKEYGLSIWKMICALSYNTETK-----MDGVRTYLM 200

Db 1076 SPSPPVTOQTSVENRRTWMNENGGKIWEGWICALTYNNDTPSGIATQIQEVRTKL 1135
QY 201 KTIYKNDIKET-----LEFASRPFLRWYEMGED 232
Db 1136 DENSKNPKLPQYKYDQVKLDDTSPDKTGSVPVSGEKITPLDFTSRPPYFRYLEWGET 1195
QY 233 FVKNRKKELVSLKKKCDSCITLRNNTSKTCD-----DNE-----NCACATQ 275
Db 1196 FCKEKKKLEKTEK-----RGDRTGHEHCSGDGYDCTRIDADRNDKFDVINCDCHIQ 1250
QY 276 CEKYKMMERMKHYSSQKKRQ-----LYKNSATYNNGLAVKANSEYTKNDPEVTAN 330
Db 1251 CRKYRKWDIKDEYHKEQEKYQGEYDKLTQDKSSGGNNCKOLEKH----- 1298
QY 331 SAKHARDYLTQLENMICTNG 351
Db 1299 -KSAAVELK--ELKHCKNG 1314

RESULT 9
T28161
hypothetical protein FC3-variant1-1 - malaria parasite (Plasmodium falciparum) (fragme
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T28161
R:Hernandez-Rivas, R.; Matfel, D.; Sterkers, Y.; Peterson, D.S.; Wellens, T.E.; Scher
Mol. Cell. Biol. 17, 604-611, 1997
A:Title: Expressed var genes are found in Plasmodium falciparum subtelomeric regions.
A:Reference number: 220483; MUID:97154495; PMID:9001213
A:Accession: T28161
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2647 <HER>
A:Cross-references: EMBL:U67959; NID:9179425; PID:91809295; PIDN:AA047438.1
A:Experimental source: strain FC027/PNG
C:Genetics:
A:Introns: 2158/3
A>Note: FC3-variant1-1

Query Match 21.1%; Score 403.5; DB 2; Length 2647;
Best Local Similarity 25.1%; Pred. No. 1e-19;
Matches 125; Conservative 60; Mismatches 131; Indels 183; Gaps 20;

QY 8 LISQIEKNNHNCCTEDAKKC-----ENKLGDEG----- 39
Db 1273 IVAGVLKKNENMCPTKRYGKAPTSMKCIPTKNTATSGSGSGALQRAKRAVES 1332
QY 40 -----EGVCMPPRRQMLCVHYLTKND-----DSKEE 66
Db 1333 GSPVTSNSGICIPRRRL--YIQKLHWASGNTVVSQAQTPQGGTSSPSGKETPSD 1389
QY 67 DLREAFIYSAAEFLROY-----NSNVNDDKIL--HRD----- 101
Db 1390 KLRTAFIOSAIEFFELMDRKKGKALAKKEKKQWVDSPLSTADPHNNVSLVIANP 1449
QY 102 -----MIPPEFRSMFYEGYRDICL-DTDLSEKIAHDVTA-KKITAVFQKI-- 150
Db 1450 NYKTCVPIPPFLRQMFYLLGDVADIFGKNDIVITPKNDKDIAREKKIKALAEKVLK 1509
QY 151 -GSKTTNGKYLEREGWMKEYGLSIWKMICALSYNTETK-----KMDGVRTYLM 199
Db 1510 MADQSPSDEK-----RQTMWQNGEHIWNGICALTYKEKDEKGTPLKQNGELKSLAMDEK 1566
QY 200 -----MKYLYKNNDIKEYLEEA-----SRPFLRWY 227
Db 1567 NKRPDKQYD--KVKLDENSGSPKIVAPAPRTTTPPPSPTSFSPRPFRYLE 1622
QY 228 EMGEDVFNKKEKELVSLKKKCDSCITLRNNTSKTCDNEN-----CGA 271
Db 1623 EMAETFCRERKKRLEKILVVC-----MDEQSKKKCGSGDGEIKKQDYSTYRDCYCE 1678
QY 272 CKQCEKYKMMERMKHHSSQKKRQLYKNSATYNNGLAVKANSEYTKNDPEVTAN 331

Db 1679 CGKCYCFYKRWIGKKKDEYDKEAYNOKTDARRNN-----NDNAFS-----TTLDPT 1726
 QY 332 AKHARDYKTOLENNICTN 350
 Db 1727 CTTAGDFLOT-LKNGPCKN 1744

RESULT 10

variant-specific surface protein 2 - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T28626
 R:Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herfelde, J.A.; Peterson, D.S.
 Cell 82, 89-100, 1995
 A:Title: The large diverse gene family var encodes proteins involved in cytoadherence an
 A:Reference number: Z20487; M0ID:95330813; PMID:7606788
 A:Accession: T28626
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2664 <SD>
 A:Cross-references: EMBL:LA0609; NID:g886376; PID:g886378; PIDN:AAA75398.1
 C:Genetics:
 A:Introns: 2197/3
 A:Note: var-2

Query Match 20.9%; Score 399.5; DB 2; Length 2664;
 Best Local Similarity 25.2%; Pred. No. 1.9e-19;
 Matches 124; Conservative 58; Mismatches 118; Indels 193; Gaps 21;

QY 1 PCKMVKLISEQIEKNNHCKTEDAK-----WKCENTK-----LGED----- 39
 Db 1281 PCQIVKELFS-----NVDTLQACSTKYGPKKNYRCRPTPTSDVYGEDGQSRVY 1334
 QY 40 -----EGYCMPPRRONLCVHYLTKIND----- 61
 Db 1335 RSTPESGNSDKKAGACIPRRRL-----YVGKLEQWANKHNTETSVSGEAREGSEAP 1391
 QY 62 -----DSKREDFREAFIKSAAEFTLL-----ROYNSKNE----- 93
 Db 1392 AREGKSSSGGKETPSDKLRTAFIESAAVETPLMDRYKKEWLAOKRAELQNGLDLYSS 1451
 QY 94 -----DDKILHRMIDPEFRSMFYEGYRDICLDTIDSEKLAHDVHTAKKKITA 145
 Db 1452 GOSDPDPNQKLLN-GVIRPDFLRLMFYTLGDYRDILVH-----GNGTSDSGNTGSMNNNT 1507
 QY 146 VROKISQKTTNGK-----KVLREB-----WKKEYGLSTWGMICALSY 184
 Db 1508 VLEASGNKEDMOKIQEKIQLIPKNGGTPLVPKSSAQTPDKMNEHAESITWKMICALTY 1567
 QY 185 -----NTEET-----KKMDE-----GVRYLAKYIKKNDIKE--- 211
 Db 1568 TEKNPDTISARGDENKTEKDEYEFKFGSTADKHGTAAPTGTGYQYDEKVKLEDTSG 1627
 QY 212 -----YLEEFASRPPLRWYTEMGEDFVNKKRELYSLKRC----- 248
 Db 1628 AKTPASASDPTPLSDVLRPPYRYLEMGONGCKRKRRIKQIKRECDMGDSKKYSGDG 1687
 QY 249 DSGTLLNNGTSNKTCDNENCGACTQCEKRYKMMERKHHSSQKKFQOLYNSATY-- 306
 Db 1688 EQCDRRD--TSNVSADLEG-RSCGNSCRFFYKWKIRKRREYDQANAYS--KQTKYE 1742
 QY 307 -NNGLAVKEANSE 318
 Db 1743 GSKAGLNDHNKE 1755

RESULT 11

B71600
 variant-specific surface protein 1 homolog PF01055c - malaria parasite (Plasmodium falciparum)
 N:Alternate names: erythrocyte membrane binding protein 1 (EMP1)
 C:Species: Plasmodium falciparum
 C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000

C:Accession: B71600
 R:Gardner, M.J.; Tetteijn, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.;
 Perera, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.
 Science 282, 1126-1132, 1998
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
 A:Reference number: A71600; M0ID:99021743; PMID:9804551
 A:Accession: B71600
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-2197 <GAP>
 A:Cross-references: GB:AE001434; GB:AE001362; NID:g38445341; PIDN:AAC71996.1; PID:g384
 A:Experimental source: clone 3D7
 C:Genetics:
 A:Gene: PF01055c

Query Match 20.8%; Score 398.5; DB 2; Length 2197;
 Best Local Similarity 25.4%; Pred. No. 1.8e-19;
 Matches 133; Conservative 61; Mismatches 132; Indels 197; Gaps 21;

QY 15 KNNIHN-----CK-----KTEDAKMKCENT-----KLGEDGVCM 45
 Db 918 KNAITNNDNLTDACKLKYCPGKEREPPNKKCVSSGKSVATAGSSGATGSKGKAICVP 977
 QY 46 PRONLCVHYLTKL-----KTEDAKMKCENT-----KLGEDGVCM 77
 Db 978 PRRRRLYVGLTFLTSAGTSSESGESSRASDVSGNGGDITTTESLRKWFJETAA 1037
 QY 78 AETFL-----ROYNSKNE-----DD-----KILHRMIDPEFR 109
 Db 1038 IETFLMHRKYEKMEQAKELQRCILLCTGASLNLGDDSNPQLOKSGTIPPLDFL 1097
 QY 110 SMETFGDYRDICLDTIDSEKIDHDV-----TTAKKRTAVPQI-----GSKPT 155
 Db 1098 LMFYTLGDYRDILRVGADDKNGNNIILNAGSNKDEKQMEKIQEKIQLIPTSGKET 1157
 QY 156 NG-----KVLREBGMKEYGLSTWGMICALSY-----NTEK 189
 Db 1158 RGPQNSVNDQSLMDLAELVHGMICALTYKDDNGKGVYKPKQIENPKIMETTK 1217
 QY 190 K-MDEGVRYLYKYYIKKNDIKE-----YLEEFASRPPLRWYTEMGEDV 234
 Db 1218 KPKDE-----KYQYQAKLEDESGEKRPDSASGKTLDFKRPYRYLEMGENIC 1270
 QY 235 KNRKKELVSLKKRCDSCITLRNG-----TSKTKDDNEN-----CGACKQCEK 281
 Db 1271 KRTEMLGRIKEDC-----YKNGRCSGDGLKCNELVIDEKIFGDLICPTCARHCRPYK 1326
 QY 282 WM-----ERMKHHSSQKKFQOLYNSATYNNGL--AVKEANSE---TYKNDPEVTEA 329
 Db 1327 WINTKRDDEFNKQSNAYSEQKKKYEENDSNQKNNGVCGTLKDDAAEFLNLRKNGPCKNES 1386
 QY 330 NSAKHARDYL-----KTQLENNICTNG 351
 Db 1387 EENKKADELDIFKRPDTEKADNCRKPCSEFKICEHNHNCSSG 1429

RESULT 12

T14602
 variant-specific surface protein - malaria parasite (Plasmodium falciparum) (fragment)
 C:Species: Plasmodium falciparum
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jun-2000
 C:Accession: T14602
 R:Voss, T.S.; Felger, I.; Weiss, N.; Beck, H.P.
 submitted to the EMBL Data Library, February 1998
 A:Description: Identification of a conserved 5' flanking region of Plasmodium falcipa
 A:Reference number: Z18158
 A:Accession: T14602
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2135 <VOS>
 A:Cross-references: EMBL:AF050740; NID:g2944094; PID:g2944095; PIDN:AAC05220.1
 C:Genetics:
 A:Gene: varph17

Query Match 20.8%; Score 397.5; DB 2; Length 2135;
 Best Local Similarity 27.5%; Pred. No. 2e-19;
 Matches 120; Conservative 51; Mismatches 149; Indels 117; Gaps 15;

10 SEQLEKNNHNCCK--TEDAKWKCENKLEDEG--VCMPPRONLCVHYLTCLND-----61
 1407 SEGSEGGSHORAKRHTSDASGE--KSAKSGEPTGSGICIPPRRRKLYVGGITKAELQSS 1465
 62 -----DKEDLEAFIKSAAEFLRLQYNNKQVND-----94
 1466 QSOALSGOTTPTAGTPOADPLLAFAVESAAVETFFLMDRYKKLNAPQSSSLGGAPLQ 1525
 95 -----DKILHRDMPPEFFRSMFYTFEGDYDIDLDT--DISEKI-ADHDVTTA 139
 1526 LKNGAISEETPEPISLKSGLTPPDLFLMFTLLGDYDIDICVGYQDYIKALEASGDKSS 1585
 140 KKKTAVFOKIGSKTTNG-----KKYLERGWMKEYGLSIWKMGCALSYNTE 187
 1586 KNPQOEISSKIEELKNGTPTPTPTVTHSPSSGTFPSSMWTNGOIHMKMICALTYKES 1645
 188 TKKMDGVR---TYLMKYI-----YKNDIKF-----YLEEFASR 219
 1646 GEKKIEOVKATDNTDLEKLEKDKYSDYDKVYEKEENETAKGPQDGLTPQTTLLSNFYKR 1705
 220 PPLRLWYTFENGEDFVKNRKEIVSLKKKC-----DSCTLRNNGTSNKTCDNE----267
 1706 PPFRLYLEEGNGNCKTRKMLKDITIECRNSNDPGHDYSGDGYHONNEKHNMMHGI 1765
 268 NCQACKTQCEKRYKWMERWKHVSQKKFOLYKNSATYNNGLAVKEANSEYKKNDEVT 327
 1766 DCRDCYECRKKRYKWKVKKRDEFHNGKNKYGEEHEKLT-----NDQNSGGGDNT 1815
 328 -----EANSAKHARDYK 340
 1816 NFOQOIKERKTAEDFLK 1832

RESULT 13

114029
 Variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
 N:Alternate names: erythrocyte membrane binding protein 1 (EMP1)
 C:Species: Plasmodium falciparum
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jun-2000
 C:Accession: T14029
 R:Chen, Q.; Barragan, A.; Fernandez, V.; Sundstrom, A.; Schlichterle, M.; Sahlen, A.; C
 J. Exp. Med. 187, 15-23, 1998
 A:Title: Identification of Plasmodium falciparum erythrocyte membrane protein 1 (PfEMP1)
 A:Reference number: Z17860; MUID:98080592; PMID:9419207
 A:Accession: T14029
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2228 <CHE>
 A:Cross-references: EMBL:AF003473; NID:q2961467; PID:q2961466; PIDN:AAC05730.1
 C:Genetics:
 A>Note: PCR3S1.2-vari

Query Match 19.6%; Score 375; DB 2; Length 2228;
 Best Local Similarity 25.9%; Pred. No. 7.8e-18;
 Matches 122; Conservative 46; Mismatches 139; Indels 164; Gaps 19;
 2 CKMVOKLISEQLEKNNHNC--TEDA-----KMKC-----ENTKLEDEGVC 43
 886 CPYVGVLT---KDN-----ESIQDACSILKYGNNRKLGRKCYTPSGEPTSSDKNCAIC 937
 44 MPPRONLCVHYLTCLND-----DKSEEDLEAFIKSA 78
 938 VPPRRRL---YIKIYDMATKTESPOASGSEASSTSGTTPDSKA-LTKAFVESAAI 993
 79 EFTLLQYNSK-----NVE-----DKILHRDMPPEFFRSMFYTFG 116
 994 ETEFLMHRKKEKKAVALAOGAGHGLPVEGSGPEYDPEDK-LKEGKIPDGLKROMFYTLG 1052

117 DYRDICL-----DTDISEK-----IAHDVTTAKKITAV-----FOKIGS 152
 1053 DYRDILFGSNDTTSVKDTPSSSDNLKINVLIAASSTDOERKMKRYKEIKNFKCST 1112
 153 KTNKKVLEREGMKKEYGLSIWKMGCALSYNTEKKNDEGYRTYLMK-----201
 1113 ERSAFNPVSHPOQWENNGKVIYHGWYCAL---TSKOKIAKGYEKKQKLENDENLMDEA 1169
 202 -----YKKNNDIKE-----YLEEFASRPPLRWTEWGEDF 233
 1170 NKKRPPOYQYTWKLDENSGISPRYTQTAQSDNPTPTLTHFYKRPYEFREWEENGESF 1229
 234 VKNRKEVLSLKKKDCSLRNNGTSNKTCDNE-----NCQACKTQCE 277
 1230 CREKKRLKIKYDCKA-----ENDQVRCGSGDEACDSITHDYSTYSPNCGCGKHC 1285
 278 KYKMKERWKHVSQKKFOLYKNSATYNNGLAVKEANSEYKKNDEVT 328
 1286 SYRKWIERKKIEFHQSNAYGQCKTDATRNNGTDFKFCCKLTETWPDAAK 1336

RESULT 14

128634
 Variant-specific surface protein 7 - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T28634
 R:Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Gulnet, F.; Herrfeldt, J.A.; Peterson,
 Cell 82, 89-100, 1995
 A:Title: The large diverse gene family var encodes proteins involved in cytoadherence
 A:Reference number: 220487; MUID:95330813; PMID:7606788
 A:Accession: T28634
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2182 <SU>
 A:Cross-references: EMBL:LA2636; NID:g886379; PID:g886380; PIDN:AAA75399.1
 C:Genetics:
 A>Note: var-7

Query Match 19.6%; Score 374.5; DB 2; Length 2182;
 Best Local Similarity 25.4%; Pred. No. 8.2e-18;
 Matches 121; Conservative 54; Mismatches 137; Indels 165; Gaps 19;
 1 PCKMVOKLISEQLEKNNHNC-----KTEDAKWKC-----ENTKLEDEGVCMPR 47
 851 PCQIVGLTFD--DKSLKEACGLKYGKGEKPPMKCYTPSGVSTATSKDAICVPPR 908
 48 RQNLGVHYL-----TKLNDSEEDLEAFIKSAAEFLRLQYNS 89
 909 RRLLYVGLSQMASRGDETFTEVSSEATSPQSSESEKLTAFIESAIIETFLMHKRYE 968
 90 KNY-----ED--DKILHRDMPPEFFRSMFYTFEDYDIDL--D 124
 969 EKKRPATQDAGLGLVSLPEPSPEGPQTOLOQTGVIPDPLRQMEYTTADYADILYSGS 1028
 125 TDISERLADHDVTTAK-----KITAVFOKIGSKTN 156
 1029 NDIISD-----TTGQTPSSSDNLKNLYLEASGSTEDEKEMKQIQAKIKIILGANS 1081
 157 GKKVLER-----EGWMEKEYGLSIWKMGCALSY-----NTETK--MDE 193
 1082 GVPPTVKNNSKYTPQWMEIADINMANVACALTYENADRGTSAKIEQKDKLKLMD 1141
 194 GVR-TYLMKYI-----KNNDIKE--YLEEFASRPPLRWTEWGEDFVKNRK 239
 1142 ANKNTPELEKYQYTNVRLKLEDESGAKSNDTIQPTLKNFVEIPTFRWLHEWGNFCFERAK 1201
 240 ELVSLKKKDCSLRNNGTSNKTCDNENC-----GACKTQCKRYKKNME 284
 1202 RLQKIRHEC-----MDGEKQYSGDGEYCEELFSKQYVNLQDLSSSCAKPCQLYTWIE 1256
 285 RMKHHVSQKKFOLYKNSATYNNGLAVKEANSEYKKNDEVT EANSAKHARDYKLT 341

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein protein search, using sw model

Run on: June 20, 2003, 14:37:29 ; Search time 15.4461 Seconds

(Without alignments)
942.514 Million cell updates/sec

Title: US-10-087-013-9

Perfect score: 1913
Sequence: 1 PCKMVKLISEQIEKNHIN.....AKHARDYKLTLENMICTNG 351

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	329	17.2	1435	EBAL_PLAFC	P19214 plasmodium
2	269	14.1	1070	PVDR_PLAVS	P22230 plasmodium
3	261	13.6	1153	PVDB_PLAKN	P50493 plasmodium
4	231	12.1	1073	PVDA_PLAKN	P22545 plasmodium
5	225	11.8	1070	PVPG_PLAKN	P50494 plasmodium
6	122.5	6.4	607	HTPG_FUSNN	O80944 fusobacteri
7	117	6.1	1391	MS72_DROHY	O08666 drosophila
8	115	6.0	944	NUP1_YEAST	P21380 saccharomyc
9	114	6.0	993	SCP1_MOUSE	O62209 mus musculu
10	110	5.8	1995	YCX7_CHIRE	P36495 chlamydomon
11	109	5.7	845	SCP1_MESAU	O60563 mesocricetu
12	106.5	5.6	896	RA50_THEAC	O90188 thermoplas
13	106	5.5	767	TOP1_CRIGR	O07050 cricetulus
14	105.5	5.5	2869	RBP1_PLAVB	O00798 plasmodium
15	104.5	5.5	817	ENPL_CARRO	P35016 catharanthu
16	103.5	5.4	2230	GOG4_HUMAN	Q13439 homo sapien
17	102.5	5.4	1976	MYHA_BOVIN	Q27991 bos taurin
18	101.5	5.3	845	VAV_HUMAN	P15498 homo sapien
19	101.5	5.3	997	SCP1_RAT	Q03410 rattus norv
20	101.5	5.3	1160	TFC3_YEAST	P34111 saccharomyc
21	101.5	5.3	1727	ALMT_SCHPO	Q04675 schistosom
22	101.5	5.3	1976	MYHA_HUMAN	P35580 homo sapien
23	101.5	5.3	2339	RPCL_PLAFA	P27625 plasmodium
24	101	5.3	1972	MYHB_HUMAN	P57749 homo sapien
25	100.5	5.3	581	Y208_BORBU	O51226 borrelia bu
26	100.5	5.3	1928	MYSL_YEAST	P08196 saccharomyc
27	100	5.2	634	NOG1_MOUSE	Q09me9 mus musculu
28	99.5	5.2	843	VAV_RAT	P54100 rattus norv
29	99	5.2	671	RPSP_HELPY	P35393 helicobacte
30	98.5	5.1	1790	USOL_YEAST	P55366 saccharomyc
31	98	5.1	443	CSN2_HUMAN	O15647 homo sapien
32	98	5.1	1539	Y373_HUMAN	O15078 homo sapien
33	98	5.1	1682	MSPL_PLAFC	P19598 plasmodium

34	98	5.1	1805	1	HMW2_MYCGE	P47460 mycoplasma
35	98	5.1	2136	1	YCF2_MARPO	P09975 methanotia
36	97.5	5.1	441	1	Y072_METJA	O60306 methanococc
37	97.5	5.1	519	1	CP7_RANDOM	O57525 rana dybows
38	97.5	5.1	726	1	HS9A_BRARE	O90474 brachydanio
39	97	5.1	609	1	YSL1_YEAST	P38280 saccharomyc
40	97	5.1	807	1	HYSA_STRAU	O59801 staphylococ
41	97	5.1	976	1	SCP1_HUMAN	O15431 homo sapien
42	97	5.1	1972	1	MYHB_MOUSE	O08638 mus musculu
43	96.5	5.0	795	1	ENPL_CHICK	P08110 gallus gall
44	96.5	5.0	845	1	VAV_MOUSE	P27870 mus musculu
45	96.5	5.0	1005	1	RA50_METJA	O58718 methanococc

ALIGNMENTS

RESULT 1	EBAL_PLAFC	STANDARD	PRT	1435 AA
AC	P19214			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	01-FEB-1996 (Rel. 33, Last annotation update)			
DE	Erythrocyte-binding antigen EBA-175.			
OS	Plasmodium falciparum (isolate Camp / Malaysia).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.			
OX	NCBI_Taxid-5835;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-90377299; Pubmed-2204835;			
RA	Sam B.K.L.;			
RT	"Sequence conservation of a functional domain of erythrocyte binding antigen 175 in Plasmodium falciparum."			
RL	Mol. Biochem. Parasitol. 41:293-296(1990).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
DR	EMBL: X52524; CAA36756.1; -			
KW	PIR: S11561; S11561.			
FT	DOMAIN 159 1104			
FT	ESSENTIAL FOR BINDING TO ERYTHROCYTES.			
FT	VARIANT 1031 1031 E -> V (IN STRAINS FCR-3 AND ITG).			
SO	SEQUENCE 1435 AA; 167389 MW; 32A4309021B1C3D6 CRC64;			
Query Match	17.28; Score 329; DB 1; Length 1435;			
Best Local Similarity	24.88; Pred. No. 2.9e-15;			
Matches	87; Conservative 63; Mismatches 113; Indels 88; Gaps 13;			
QY	14 EKNNHCKCTEDAKKCKCKTGLGDEGVCMPPRRQVLCVHYLTKLNDSDS-----KEEDLR 69			
DB	463 DKNSV-----DTNFKWCKKMPYILSTDVCPVRROELCGNIDRLTKMLIKETILA 518			
QY	70 EAFKSAAEFTFLROYNKRNVEDDKILRDMIPPEFRSMFTFGDYRDICLDPISE 129			
DB	519 IAIYES-----RLKRRKKNKDKEVCILINK-----TFNDIDIDIGGTQYWN 561			
QY	130 KIDHDVYTTAKKKTITAVFOKIGSTTKGKV-----LEREGWKEGLSTWKMLCA 181			
DB	562 DLSNRKL-----VGKINTSKKYVHRKKNKDKLEFRDWMVVIKKDVVN----- 603			
QY	182 LSVTEFKMKDEGVRTYLMKYIYKNNDI-KEYLEEFASRPFLRWVEMGEDFVKNRKE 240			
DB	604 -----VISVFEKDKYVCE--DDINIRQFFRWSEKGDVYQDQTKM 644			
QY	241 LVSLKKKDCSTLRNNGTSNKTCDNENGCACKTQCEKRYKWRKWKHYSSQKKRPOLY 300			

Db 645 IEFLEKCEC-----KEKPEEDN-----CKSKCNSEYKEMISKEEYNQAKQY 690
 QY 301 KNSATNNGLAVEANSEY--KNDEVEANSARFARYLTQLEMTCT 349
 Db 691 QKGNNTKMSSEFSTPEYLLKYSEKCSNLNFEDEFKEELHSDYKN-CT 740

RESULT 2

PVDR_PLAVS
 ID PVDR_PLAVS STANDARD: PRT: 1070 AA.
 AC P22290:
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Duffy receptor precursor (Erythrocyte binding protein).
 GN PVDR.
 OS Plasmodium vivax (strain Salvador I).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=126793;
 (1)

SEQUENCE FROM N.A.

MEDLINE=91187056; PubMed=1849231;
 RA Fang X., Kaslow D.C., Adams J.H., Miller L.H.;
 RT "Cloning of the Plasmodium vivax Duffy receptor."
 RL Mol. Biochem. Parasitol. 44:125-132(1991).
 CC - FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP
 CC DETERMINANT.
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - SIMILARITY: HIGH, TO P. KNOWLESI DUFFY RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC EMBL: M61095; AAA63423.1; -

DR Malaria; Receptor; Glycoprotein; Signal; Transmembrane.
 KW SIGNAL 1 20
 FT CHAIN 21 1070 POTENTIAL.
 FT DOMAIN 21 1007 DUFFY RECEPTOR.
 FT TRANSMEM 1008 1025 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1026 1070 POTENTIAL.
 FT CARBOHYD 183 183 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 715 715 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 787 787 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 825 825 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 903 903 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 938 938 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 1070 AA; 119683 MW; CB051DF13E294603 CRC64;

Query Match 14.1%; Score 269; DB 1; Length 1070;
 Best Local Similarity 26.2%; Pred. No. 3e-11;

Matches 84; Conservative 50; Mismatches 111; Indels 76; Gaps 15;

QY 6 OKLISEOI-----EKNNNHNC--KTEPDAKKKCENTKLGDEGVCMPPRONLCVHYL 56
 Db 196 KRLISSALINHAFLONTVMKNCNKRKRERDMDC-NFK--KDCVTPDRRYQLCKWKL 250
 QY 57 TKL--NDSS--KEEDIREAFIK-----SAAETFLRQYNSKNVEDDKILHRDMLPE 106
 Db 251 TNLVNNNDJFHFHDIKRLKLYDAVAGDLLLKNNRYKN----- 298

QY 107 FFFSMFTFGDYDIDICDIDIS-----EKIDHDTYTTAKKKTIVFQKISGTTNGKVVLE 162
 Db 299 FCDIRMSLGDGFDIDIGTIDMIGISGYKVEEN-----LRSIF-----GTDERAQR 345

QY 163 REGWKEKGLSTWKMALCALSTNTETKMDGCVRTYLLKYYTKNNDIKYLEEFASRP 222
 Db 346 RQWMESEKRAQIWTAMMSV-----KKRLGNFWICKLVAVN-----IEPQI 389
 QY 223 LRWVTEWGEDEFYKNNKRELVSLLKNGSTSKCTDDDNENGACACKTOCEKRYKW 282
 Db 390 YRMIREMKRDYSELPTFVOKLKEKCG--KINTDKVC-----KVPCCNACKSYDOW 442

QY 283 MERWKKHYSQKKRFRQLYKNS 303
 Db 443 ITRKKNQMDVLSNKEFISYKNA 463

RESULT 3

PVDR_PLAVS
 ID PVDR_PLAVS STANDARD: PRT: 1153 AA.
 AC P50493:
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Duffy receptor, beta form precursor (Erythrocyte binding protein).
 OS Plasmodium knowlesi.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5850;
 (1)

SEQUENCE FROM N.A.

MEDLINE=92357776; PubMed=1496004;
 RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
 RT "A family of erythrocyte binding proteins of malaria parasites."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992).
 CC - FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP
 CC DETERMINANT.
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - SIMILARITY: HIGH, TO P. VIVAX DUFFY RECEPTOR.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC EMBL: M90694; AAA29603.1; -

DR Malaria; Receptor; Glycoprotein; Signal; Transmembrane;
 KW Multigene family.
 FT CHAIN 1 21
 FT SIGNAL 22 1153 POTENTIAL.
 FT DOMAIN 22 1085 DUFFY RECEPTOR, BETA FORM.
 FT TRANSMEM 1086 1106 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1107 1153 POTENTIAL.
 FT TRANSMEM 1107 1153 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 348 348 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 467 467 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 626 626 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 722 722 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 847 847 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 856 856 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 900 900 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 910 910 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 935 935 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 1153 AA; 130471 MW; 6497BD16CE7BEC01 CRC64;

Query Match 13.6%; Score 261; DB 1; Length 1153;
 Best Local Similarity 23.9%; Pred. No. 1.2e-10;
 Matches 87; Conservative 50; Mismatches 125; Indels 102; Gaps 14;

QY 3 KWOXKLISBO--IEKNNNHNC--KTEPDAKKKCENTKLGDEGVCMPPRONLCVHYLTK 58

```

Db      195  RMCSGVINOTFLQKNVRRRCNKRKRGTRDWCPTPK-----DVCIPDRRYQLCMKELTN 249
      59  LNDSSKEE-----DLREAFIKSAAEFFLLQYNSKNVNEEDKILHRMIPPEFF 108
      250  LVNNTTSHNDITFLKLNKEKLYDAVEGDLTKKNYNSSED-----LC 297
      109  RSMFEYGDYRDICLDTDISEKIAHDVYTTAKKITAFAVOKISGKTTNGKKVLEREGWKK 168
      298  KOIKMSLEDFGDIIMGTDM-EGIGYSQV--VENNLTFV-----GTGKTOLDRRKKWN 348
      169  EYGLSIWKMALCALSYNTETKKMDGVRYYLKKYIKKNDIKYLEEFAFRPPLRWYTE 228
      349  ESKKYITWEATILISV-----KKKLNQYSAMNCK-----EDVOINVE-----POIYRMIRE 392
      229  WGEDEFYKNNKKELYSIKKKCD-----CTLRNNGTSNKTCDNENCGACKTCCEKY 279
      393  WGDVYSELPRKQEKKEKCDKRLYYTNLRICTM-----SPCNDCKLY 436
      280  KKMERNKKHYSSQKKFQLYKNSATYNNGLAVKEANSEYTKNDPEVTANSKAKHARDYL 339
      437  DQWITKKKQWDLSTKFSYVKKG-----QIETENTITAYDIL 475
      340  KTDL 343
      476  KOEL 479

Db      476  KOEL 479

RESULT 4
PVDA_PLAKN      STANDARD:      PRT: 1073 AA.
AC  P22545;
DT  01-AUG-1991 (Rel. 19, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  01-OCT-1996 (Rel. 34, Last annotation update)
DE  Duffy receptor, alpha form precursor (Erythrocyte binding protein).
OS  Plasmodium knowlesi.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5850;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  MEDLINE=92357776; PubMed=1496004;
RA  Adams J.H., Hudson D.E., Torii M., Ward G.E., Wellens T.E.,
RT  Alkawa M., Miller L.H.;
RT  "The Duffy receptor family of Plasmodium knowlesi is located within
RT  the micromeres of invasive malaria merozoites.";
RL  Cell 63:141-153(1990).
CC  -!- FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP
CC  DETERMINANT.
CC  -!- SIMILARITY: HIGH, TO P.VIVAX DUFFY RECEPTOR.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: M90466; AAA29602.1; -
CC  EMBL: M68517; AAA29590.1; -
CC  EMBL: M68518; AAA29591.1; -
CC  PIR: A35970; A35970.
CC  Malaria; Receptor; Glycoprotein; Signal; Transmembrane;
CC  Multigene family.
CC  SIGNAL 1 21 POTENTIAL.

```

```

FT      CHAIN 22 1073 DUFFY RECEPTOR, ALPHA FORM.
FT      DOMAIN 22 1007 EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM 1008 1029 POTENTIAL.
FT      DOMAIN 1030 1073 CYTOPLASMIC (POTENTIAL).
FT      CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 348 348 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 746 746 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 779 779 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE 1073 AA; 120683 MM; 3965FC9F46B71808 CRC64;

Query Match 12.18; Score 231; DB 1; Length 1073;
Best Local Similarity 22.88; Pred. No. 1,3e-08;
Matches 86; Conservative 59; Mismatches 146; Indels 86; Gaps 15;

      3  KMYQKLISEQIEKNNI-----HNCKRTEDAKKCENTKLGDEGVCMPPRONICVHYLTK 58
      195  RMSNGVINOTFLQNNVMDKCNDRKRGTRDWCPTPK-----AEKICISVRRYQLCMKELTN 249
      59  LNDSSKEEDIRE-APIKSAAEFFLLQYNSKNVNEEDKILHRD--MIPPEFRSMFYTF 115
      250  LVNNTTSHNDITFLK-----LNLKRKLIMYDAVEGDLTKKNYQYNEFCQDIRWGL 304
      116  GDYRDICLDTDISEKIAHDVYTTAKKITAFAVOKISGKTTNGKKVLEREGWKEYGLSTW 175
      305  GDFGDIIMGTNM-EGIGYSQV--VENNLROVF-----GTDEKAKODRKQMNESKEHW 355
      176  KGMICALSYNTETKKMDGVRYYLKKYIKKNDIKYLEEFAFRPPLRWYTEMGEDFVK 235
      356  RAMMFSIRSLRKEK-----FWICK-KVITLKE-----POIYRIRRMGNDYMS 399
      236  NRKKEVLSLKKKDCSCTLRNNGTSNKTCDNENCGACKTCCEKYKKMERNKKHYSSQKK 295
      400  KLRKEQGLKNEKCAKSLYYNN--MAIC-----MLPICHDAKSDYITKKKQWDLST 452
      296  KFOLYKNS-----ATYNNGL-----AVKEANSE 318
      453  KFSYVKKTKQKIGTENIATAYDILKQELNGKREAFENEIKRKNLYNHLCPVVEARKN 512
      319  TYKNDPEVTANSKAKHA 335
      513  TQENYKVNQSGVESKAA 529

Db      513  TQENYKVNQSGVESKAA 529

RESULT 5
PVDA_PLAKN      STANDARD:      PRT: 1070 AA.
AC  P50494;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  01-OCT-1996 (Rel. 34, Last annotation update)
DE  Duffy receptor, gamma form precursor (Erythrocyte binding protein).
OS  Plasmodium knowlesi.
OC  Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX  NCBI_TaxID=5850;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  MEDLINE=92357776; PubMed=1496004;
RA  Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
RT  "A family of erythrocyte binding proteins of malaria parasites.";
RT  Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992).
CC  -!- FUNCTION: BINDS TO THE HUMAN ERYTHROCYTES DUFFY BLOOD GROUP
CC  DETERMINANT.
CC  -!- SIMILARITY: HIGH, TO P.VIVAX DUFFY RECEPTOR.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC  the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@lsb-sib.ch).

DR EMBL: M90695; AAA29604.1; -
KW Malattia; Receptor; Glycoprotein; Signal; Transmembrane;
FT Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 1070
FT TRANSMEM 22 1003
FT TRANSMEM 1004 1025
FT DOMAIN 1026 1070
FT CARBOHYD 134 134
FT CARBOHYD 179 179
FT CARBOHYD 676 676
FT CARBOHYD 743 743
FT CARBOHYD 785 785
FT CARBOHYD 936 936
FT SITE 279 281
SEQUENCE 1070 AA; 120931 MW; 703068811BC11B50 CRC64;

POTENTIAL.
DUFFY RECEPTOR, GAMMA FORM.
EXTRACELLULAR (POTENTIAL).
POTENTIAL.
CYTOPLASMIC (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
CELL ATTACHMENT SITE (POTENTIAL).
Query Match 11.8%; Score 225; DB 1; Length 1070;
Best Local Similarity 20.3%; Pred. No. 3.3e-08;
Matches 80; Conservative 53; Mismatches 135; Indels 126; Gaps 13;

QY 3 KAVOKLISEQLEKNTI---HNCKTEDAKWKCENTLGEDEGCMPPRQNLCAVHLTK 58
DB 195 KMSGAINQIFLQNNVNDKCDKRRKGEKRDWDCPTK---DVCIDRRYQLCMEITN 249
QY 59 LINDS-----KEEDLREAFIKSAAEFTFLROYNSKNVDDTLHRDMIPPEFR 109
DB 250 LVDTDPHSHDIIRKYSRRLLIYDVGKRDLLKRYNNVSD-----LCK 297
QY 110 SMFTFSDYDIDCDLTIDISEKIDADHYTAKKITYAVQKIGSKTNGKVLREGWKE 169
DB 298 DIKSLDQDFGDIIGDTM-EGIGSLV--VONNLRSIF-----GTGTSALDRKKWMD 348
QY 170 YGLSIWKGMLCALSS---YNTETKKMGEGVRYTIKTYKNNIDKEVLEFASAPPLRW 225
DB 349 HKKIDIMKAMILSVKEKNTYSAMNCKEDVQIKV-----EPQYRW 387
QY 226 VTEGDEFEVKNRKEKELYSKKKCDSS-----CTLRNNGTSNKTCDNENGCACKTQC 276
DB 388 IRENGRDYSEFRQRRKLNKCEKDKLYSTMLICTL-----PPCNMAC 431
QY 277 EYKKMKMRKKKHHSSQKKFQYK-----NSATYNNGL-- 310
DB 432 KSYDEMTIGKKRQMDVLSTKFSYVKKAKIETENIARAYDILKQELNGFNEVTEENIK 491
DB 311 -----AVKEANSEYTKNDPEVTEANSK 333
DB 492 RDKLYNFCVCIQDEAKRNTQENKKNVSGVEAK 525

RESULT 6
HTPG_FUSNN STANDARD: PRT: 607 AA.
AC Q08GH4, 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chaperone protein htpg (Heat shock protein htpg) (High temperature protein G)
GN HTPG OR FMO321.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 25586;
RX MEDLINE=2186934; PubMed=11889109;
RA Kaptirai V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,

RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vastisa O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,
RA Larssen N., D'Souza M., Malinas T., Pusch G., Haselkorn R.,
RA Fornslein M., Kyrides N., Overbeek R.,
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
RT J. Bacteriol. 184:2005-2018(2002).
CC - FUNCTION: Molecular chaperone. Has ATPase activity (By
CC similarity).
CC - SUBUNIT: Homodimer (By similarity).
CC - SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC - SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@lsb-sib.ch).

DR EMBL: A6010544; AA194527.1; -
DR PROSITE: PS00298; HSP90; 1.
KW Chaperone; ATP-binding; Heat shock; Complete proteome.
FT DOMAIN 1 323 A: SUBSTRATE-BINDING (BY SIMILARITY).
FT DOMAIN 324 534 B (BY SIMILARITY).
FT DOMAIN 535 607 C.
SQ SEQUENCE 607 AA; 70368 MW; D1519D82B0FD1B08 CRC64;

Query Match 6.4%; Score 122.5; DB 1; Length 607;
Best Local Similarity 24.1%; Pred. No. 0.21;
Matches 66; Conservative 48; Mismatches 121; Indels 39; Gaps 11;

QY 101 DMIPPEFRMFTFC---DYRDICLDIDISKINDHVTY--AKKITYAVFQKIGST 154
DB 299 DLI-PEYFN--FISGVDDSLNSIRLLQDQNAELQYISNLEKTTSEKTL--LK 352
QY 155 TNGKVLREGWKEKELYSKKKCDSS-----EGVRYLMKTYKN 206
DB 353 NDREKYE--FWKEGSRISKAGVDMFGNKR-KLQDILLIFVSHDDKYTLKEVDIM 408
QY 207 NDIKEYLEFASAPPLRWTEGDEPVNRRKELYSKKKCDSCILRN----- 255
DB 409 GDNKEILYPAESVDARAKYLPK--EKLQEGEVLLIDKIDFTLMVARDYSGKEFS 466
QY 256 -NGTSNKTCDNENGCACKTQCEKRYKMKRKHYSQKKFQYKNSATYNNGLAVE 314
DB 467 INSDFEFSDKKEKEEYVKIADENKELIEKAEFLKDKYSEVLSNNIGNSASSLLAKG 526
QY 315 ANS-ETVKNPEVTEANSKAKHARDYLTQLENNI 347
DB 527 GLSLEMEKTLSEMTNNNDMPKAEKVALAINPEHVL 560

RESULT 7
MST2_DROHY STANDARD: PRT: 1391 AA.
AC Q08696;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Axoneme-associated protein mst101(2).
GN MST101(2).
OS Drosophila hydei (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7224;
RN [1]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
RC MEDLINE=9504538; PubMed=7957199;
RX Neesen J., Padmanabhan S., Buenemann H.;
RA "Randomly arranged repeats of a novel highly charged 16-amino-acid

```
DT      16-0crt-2001 (Rel. 40, Last annotation update)
DE     NUPI protein (Spindle poly body spacer protein Spcl10).
GN     NUPI OR SPCL10 OR YDR356W OR D9476.3.
OS     Saccharomyces cerevisiae (Baker's yeast).
OC     Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC     Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX     NCBI_TaxID=4932;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=S288C;
RA      Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA      Favetto A., Fulton L., Gattung S., Greco T., Kirsten J.,
RA      Kubaca T., Hallsworth K., Hawkins V., Hillier L., Jier M.,
RA      Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA      Mardis E., Meneses S., Miller N., Nhan M., Pauley A., Peluso D.,
RA      Rifken L., Riley L., Taich A., Trevisan E., Vignati D.,
RA      Wilcox L., Wolldman P., Vaudin M., Wilson R., Waterston R.;
RL      Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC      -1 FUNCTION: ESSENTIAL COMPONENT OF THE NUCLEOSKELETON. POTENTIAL
CC      ROLE IN CROSSLINKING FILAMENTS OR ANCHORING OTHER MOLECULES. IT
CC      IS ESSENTIAL FOR GROWTH.
CC      -1 SUBCELLULAR LOCATION: NUCLEAR; TIGHTLY ASSOCIATED WITH THE
CC      NUCLEUS. IT IS PRESENT IN A GRANULAR PATTERN THAT EXCLUDES THE
CC      NUCLEOLUS.
CC      -1 PTM: MAY BE REGULATED BY PHOSPHORYLATION EVENTS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X11582; CAAT7668.1; -
DR      EMBL; X73297; CAA51733.1; -
DR      EMBL; U28372; AAB64791.1; -
DR      PIR; S26710; S26710.
DR      PIR; S34288; S34288.
DR      SGD; S0002764; NUPI.
KW      Coiled coil; Nuclear protein; Phosphorylation.
FT      DOMAIN          164       791       COILED COIL.
FT      DOMAIN          54        59        NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT      DOMAIN          726       731       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT      DOMAIN          742       747       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT      DOMAIN          731        944       ARG/TYR-RICH.
SO      SEQUENCE         _944 AA; 111781 MW; 04FAA074BB8A0BC8 CRC64;
OY      Query Match      6.0%; Score 115; DB 1; Length 944;
OY      Best Local Similarity 18.9%; Pred. No. 1,2;
OY      Matches 79; Conservative 64; Mismatches 132; Indels 142; Gaps 16;
DB      15 KNHINCKTEDAKWKCENTKAGEDGVCMPPRRQ-----LRAPFISAA 77
DB      197 NKRLQALEKEUDAKNKKIYNRSKRVDHSCIEERRGMERKLAELEBKLTQVQDYLEENN 256
OY      50 -----NLCVHYITLKRLNDSGSEED-----LRAPFISAA 77
```

```

Db 257 SDVQSLKLSKDEDELNL-MNELNELKSNAEKQLEFKNELKRNKELNELKIKSD 315
OY 78 AETFLROYNSKNVEDDILHRDMLPPEFFSMFTFDYDIDLDITDISE-----KIA 132
Db 316 MDLQLOKONKESRIKDE--LNE-----LETFSENGSOSAK 351
OY 133 DHDVTTAKKITAFAVOKISGTTNGKVLREGWMEKGLSIWGMICALSYNTEKMD 192
Db 352 ENEMLKMLKKIALEEEISTKNS--QLAKEG-----KLASLMQLDLEKLNROSO 403
OY 193 EGVRYLYMYIKNN-----DIKEYIEEASAPPLRWVTEWGEDFVKNRKELVSLKK 247
Db 404 LGSREELK--KTNDKLOKODRIAREEVS-----KDERIIDLOK 442
OY 248 C-----DSCTLRNNGTSNKTCDN-----ENCGACKTOCEKYMERMK-KH 289
Db 443 VKOLENDLFIKTKHSEKSTITDNELESKDLIKILLEN--DLKVAQEKYSKMEKELKE 500
OY 290 YSSQKKKPOLYKNSATYNNGLAVKEANSETYKNDPEVTANSKARDYLKTOLENM 346
501 FVYKISEKLEDEKTTLMNEKISNLAEANSOLKNKIE-DNSTATHMKENYEQLESL 556

```

RESULT 9

SCPL_MOUSE

ID SCPL_MOUSE STANDARD: PRT: 993 AA.

AC 062209: 062329: 009205: P70192:

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Synaptonemal complex protein 1 (SCP-1 protein).

GN SCPL OR SCPL

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-CBA; TISSUE-Testis;

RX MEDLINE=96004899; PubMed=7548215;

RA Sage J., Martin L., Guzin F., Rassoulzadegan M.;

RL "cDNA sequence of the murine synaptonemal complex protein 1 (SCP1).";

RL Biochim. Biophys. Acta 1263:258-260(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-Swiss; TISSUE-Testis;

RX Kerr S.M., Taggart M.H., Lee M., Cooke H.J.;

RL Submitted (APR-1995) to the EMBL/Genbank/DBJ databases.

RN [3]

RP SEQUENCE OF 1-149 FROM N.A.

RA STRAIN-C57BL/6; Sage J., Li Y., Martin L., Mattel M.-G., Guenet J.-L., Liu J.G.,

RL Hoog C., Guzin F., Rassoulzadegan M.;

RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.

RN [4]

RP SEQUENCE OF 95-787 FROM N.A.

RA STRAIN-ICR; TISSUE-Testis; Tsuchida J., Nishina Y., Nozaki M., Uchida K., Nishimune Y.;

RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: MAJOR COMPONENT OF THE TRANSVERSE FILAMENTS OF

CC SYNAPTONEMAL COMPLEXES (SCS), FORMED BETWEEN HOMOLOGOUS

CC CHROMOSOMES DURING MEIOTIC PROPHASE.

CC -1- SUBCELLULAR LOCATION: NUCLEAR. IN TRIPARTITE SEGMENTS OF

CC SYNAPTONEMAL COMPLEXES, BETWEEN LATERAL ELEMENTS IN THE NUCLEUS.

CC FOUND ONLY WHERE THE CHROMOSOME CORES ARE SYNAPSED. ITS N-TERMINUS

CC C-TERMINUS EXTENDS WELL INTO THE SYNAPTONEMAL COMPLEX WHILE THE

CC SYNAPTONEMAL COMPLEX (BY SIMILARITY).

CC -1- DOMAIN: CONSISTS OF AN ALPHA-HELICAL STRETCH OF 700 AA RESIDUES,

CC FLANKED BY N- AND C-TERMINAL GLOBULAR DOMAINS. THE C-TERMINAL

CC DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

DR EMBL: Z38118; CAAB6262.1;
DR EMBL: L41069; AAB64514.1; ALT. INIT.
DR EMBL: U62864; AAC53335.1; JOINED.
DR EMBL: U62860; AAC53335.1; JOINED.
DR EMBL: U62861; AAC53335.1; JOINED.
DR EMBL: U62862; AAC53335.1; JOINED.
DR EMBL: U62863; AAC53335.1; JOINED.
DR EMBL: D86539; BAA13639.1;
DR MGD: MGI:105931; Sycp1.
KW Nuclear protein; Meiosis; Cell division; Phosphorylation;
KW DNA-binding; Coiled coil.
FT DOMAIN 12 97 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 104 815 COILED COIL (POTENTIAL).
FT DOMAIN 114 117 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 697 700 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 898 901 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 978 986 ARG/LYS-RICH (BASIC).
FT CONFLICT 527 527 F -> L (IN REF. 2).
SQ SEQUENCE 993 AA; 115962 MW; 1A4FA790D64FAFE6 CRC64;

```

Query Match 6.08; Score 114; DB 1; Length 993.

Best Local Similarity 19.5%; Pred. No. 1.4;

Matches 82; Conservative 63; Mismatches 163; Indels 112; Gaps 15;

```

OY 2 CKWYQXLSIPQIEKNH-----NCKKT-----EDAKKCEBN-----K 35
Db 180 CNLKEVCASAEKTNKYEERETROYVYDLSNIEKMLAEELRVAQENARLEWHF 239
OY 36 LGEDEGVCMPRRONLCVHYLTFLKINDSKEEDIREFAFKSAAE-----TFLLROYNS 89
Db 240 LKEDH-----EKIQHLEERYQKEVN--NKNQVSELLQSAEENKMKDLTFLEESRDK 292
OY 90 KNVEDDKILHRDMLPPEFFSMFTFDYDIDLDITDIS--EKIADHDVTTAKKIT-- 144
Db 293 ANOLEEKTQLDNDLSELSEKDHITSELEDIMQSRMSTQKALEEDILQIATKTISQI 352
OY 145 -----AVFKRIG--SKTNGKVLREGWMEKGLSIWGMICALS--YNTETPKMDEG-- 194
Db 353 TEVKEMQELNLAKTHSVTE-----LKATTCYLEELREDOORLEKNEED 400
OY 195 -VRYLYMYIKYKNDIKLEYLEFASRPPLRWVTEWGEDFVKNRKELVSLK----- 245
Db 401 QLKLTIVELQKNSNLEEMTK-----FKNNKEVELEIKNILLAEODK 442
OY 246 -----KKCDSCFLRNNGTSNKTCDNENCGACKTOCEKYMERMK-KH 289
Db 443 LDEKROVEKLAELQEKDELFTLETREREKVDHLDQEVYTVKISEOHLKOVEREMKTE 502
OY 290 YSSQKKKPOLYKNSA--TYNGLAVKEANSETYKNDPEVTANSKARDYLKTOLENM 346
Db 503 LEKELKLNELTASCMILLKMKFQGEASDMLAEIKKHODILINCKQKEEFLMLQIENL 562

```

RESULT 10

YCX7_CHLRE

ID YCX7_CHLRE STANDARD: PRT: 1995 AA.

AC P36495: 037303:

DT 01-JUN-1994 (Rel. 29, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 232.2 kDa protein (ORF1995) (ORF-S) (ORFA).

OS Chlamydomonas reinhardtii.

OG Chloroplast.

OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

OC Chlamydomonadaceae; Chlamydomonas.

OX NCBI_TaxID=3055;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97218038; PubMed-9065699;
RA Boudreau E., Turmel M., Goldschmidt-Clermont M., Rochaix J.-D.,
RA Sivan S., Michaels A., Ieu S.,
RA "A large open reading frame (orf1995) in the chloroplast DNA of
RT Chlamydomonas reinhardtii encodes an essential protein.";
RT Mol. Gen. Genet. 253:649-653(1997).
RN [2]
RP SEQUENCE OF 1573-1995 FROM N.A.
RC STRAIN-CW15;
RA Sivan S., Michaels A.;
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1925-1995 FROM N.A.
RX MEDLINE-87031585; PubMed-2876928;
RA Moessner J.P., Gilham N.W., Boynton J.E.;
RA "The sequence of the chloroplast atpB gene and its flanking regions
RT in Chlamydomonas reinhardtii.";
RT Gene 44:17-28(1986).
RN [4]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE-94207185; PubMed-8155879;
RA Boudreau E., Olis C., Turmel M.;
RA "Conserved gene clusters in the highly rearranged chloroplast genomes
RT of Chlamydomonas moewusii and Chlamydomonas reinhardtii.";
RT Plant Mol. Biol. 24:585-602(1994).
CC -1- FUNCTION: ESSENTIAL FOR CELL GROWTH. MAY BE INVOLVED IN BINDING
CC CHLOROPLAST DNA TO EITHER THE CHLOROPLAST ENVELOPE OR THE
CC THYLAKOID MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO C.VULGARIS ORF819.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X82726; CAA63385.1; -
DR EMBL: X76934; CAA54257.1; -
DR PIR: S41289; S41289.
DR InterPro: IPR001123; Lyse.
DR Chloplast; Transmembrane; Thylakoid.
FT TRANSMEM 31 51 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 212 232 POTENTIAL.
FT TRANSMEM 254 274 POTENTIAL.
FT TRANSMEM 307 327 POTENTIAL.
FT TRANSMEM 1588 1588 K->R (IN REF. 2).
FT CONFLICT 1588 1588
SO SEQUENCE 1995 AA; 232194 MW; E5A3F167EAF5FC8A CRC64;

Query Match
Best Local Similarity 19.7%; Score 110; DB 1; Length 1995;
Matches 92; Conservative 61; Mismatches 171; Indels 142; Gaps 19;

```

```

DB 1214 KSLVYNOTLRLKVTBDKQDMWRTKORVITKRKSARKRDRFRKQIAYVKKLLAALSKEVE 1273
DB 188 TKKM-----DEGRITLMK-----YIYNNDIKVELEFASRPPRLRAY 226
DB 1274 TEKSNLYOTLYGVEYISDYLRLRVPTGSSAVIDSTVLRKKODQAOLPEETNNVQNSFV 1333
DB 227 TE-----WGEDFYVNRKRELVSLLKKDCSCTL-----RNGTSNKTCDNENC--G 270
DB 1334 DSNNNWVQTFPAKKLKKKISSKGRYRSLSLARYLATATRRPRLYGDLNLTIKDITTTLOG 1393
DB 271 ACTQCEKY-----KKWERKKHYSSQ-KKKPOLYKNSATY 306
DB 1394 AFITKEKODSLNLTQKROELNSLKSQIKKRSRHSWKRRHQFSRHHYKRRKHTH 1453
DB 307 NNG-LAVKEANSETTYKNDEPVEYEAANSARIAADYLT-OLENMTCTN 350
DB 1454 GNGKLRVMKKLKKFRATNELRWMMWSFLPRYLSMLQVNNSTLTN 1499

RESULT 11
SCPL_MESAU
ID SCPL_MESAU STANDARD; PRT; 845 AA.
AC Q60563;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Synaptonemal complex protein 1 (SCP-1 protein) (Meiotic chromosome
DE synapleic protein) (Fragment).
GN SCP1 OR SYN1.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95181577; PubMed-7876343;
RA Dobson M.J., Pearlman R.E., Karaiskakis A., Spyropoulos B.,
RA Moens P.B.;
RA "Synaptonemal complex proteins: occurrence, epitope mapping and
RT chromosome disjunction.";
RL J. Cell Sci. 107:2749-2760(1994).
CC -1- FUNCTION: MAJOR COMPONENT OF THE TRANSVERSE FILAMENTS OF
CC SYNAPTONEMAL COMPLEXES (SCS), FORMED BETWEEN HOMOLOGOUS
CC CHROMOSOMES DURING MEIOTIC PROPHASE (BY SIMILARITY). HAS NON-
CC SPECIFIC DNA BINDING CAPABILITY.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. IN TRIPARTITE SEGMENTS OF
CC SYNAPTONEMAL COMPLEXES, BETWEEN LATERAL ELEMENTS IN THE NUCLEUS.
CC FOUND ONLY WHERE THE CHROMOSOME CORES ARE SYNAPSED. ITS N-TERMINUS
CC C-TERMINUS EXTENDS WELL INTO THE LATERAL DOMAIN OF THE
CC SYNAPTONEMAL COMPLEX.
CC -1- DOMAIN: CONSISTS OF AN ALPHA-HELICAL STRETCH OF 700 AA RESIDUES,
CC FLANKED BY N- AND C-TERMINAL GLOBULAR DOMAINS. THE C-TERMINAL
CC DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L32978; AAC42039.1; -
DR Nuclear protein; Meiosis; Cell division; Phosphorylation;
KM DNA-binding; Coiled coil.
FT NON_TER 1 1
FT DOMAIN <1 672 COILED COIL (POTENTIAL).
FT DOMAIN 553 556 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 753 756 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

```


FT DOMAIN 830 838 ARC/LYS-RICH (BASIC).
 SQ SEQUENCE 845 AA; 99401 MW; D7F28873C824C6A8 CRC64;
 Query Match 5.7%; Score 109; DB 1; Length 845;
 Best Local Similarity 20.0%; Pred. No. 2.7;
 Matches 82; Conservative 57; Mismatches 179; Indels 92; Gaps 15;

QY 2 CKWVKLISQIEKNNH-----NCKKT-----EDAKMKCENT-----K 35
 Db 36 CNLKETSARSAAKTKKYEERETROYVDLNNNEKMLAEELRVQAMNRLDMHFK 95
 QY 36 LGEDEVCPMPRRONLCVHYLTKLNDKSKEEDLREAFISAAE-----TELLROYNS 89
 Db 96 LKEDH-----EKIOHQEEYKKEVND--KENQVSLTIQTEKENMKKDLTFLEESRDK 148
 QY 90 KKNEDKILHRMDIPPEFFSMFTYFGDYRIDCLDDIS---EKIDHVTYTKKITYL 146
 Db 149 VNQLEKTKLQDENVVELNKKDKHLTSEEDTKMSLQSRNNTQKALEEDQITKITYOL 208
 QY 147 FOKIGSKTNGKKVLEPREGMKKEYLSIMKGMCLASYNTEKKMGVRYTLMKITYNN 206
 Db 209 TE-----EKEAMEEFPNKAKTIDHSFMVTELKATCTLEELTTEQORLYKNE 255
 QY 207 NDIKEYLEEFASPPPLRWVTEWGEDFVKNRKKEVLSKK----- 246
 Db 256 DQKILTMELQKSNLEDEWTK---FKNNNEVKLEELKILAEQDKLDEKKQVEKLA 311
 QY 247 ----KDCSCLRNNGSNKTCDDNENCGACKTCCEKTKKMKEMKKHYSSQ-KKPOLYK 301
 Db 312 ELQGGKDELTLLQTRKEVHDLDEQLLVKISDQNSKQVEELTKLEBEKLNMAELTA 371
 QY 302 N--SATYNNGLAVKEANS--ETYYKNDPEVTEANSAAKHADYLTQLENN 346
 Db 372 SCCKLSIENKKLQETNDMALELKXQEDIT--NSKKQOEEMLK-QLENN 418

RESULT 12
 ID RASO_THEAC STANDARD; PRT; 896 AA.
 AC 09HTR8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA double-strand break repair rad50 ATPase.
 GN RAD50 OR TA0157.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
 OC Thermoplasmataceae; Thermoplasma.
 NCBI_TaxID=2303;
 [1]
 SEQUENCE FROM N.A.
 RX STRAIN-DSM 1728;
 RX MEDLINE=20479972; PubMed=11029001;
 RA Repp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
 RT acidophilum.";
 RT accession: 407:508-513(2000).
 RT Nature 407:508-513(2000).
 CC -FUNCTION: Involved in DNA double-strand break repair (DSBR). The
 CC rad50/mre11 complex possesses single-strand endonuclease activity
 CC and ATP-dependent double-strand-specific exonuclease activity.
 CC Rad50 provides an ATP-dependent control of mre11 by unwinding
 CC and/or repositioning DNA ends into the mre11 active site (By
 CC similarity).
 CC -SUBUNIT: Forms a complex with mre11 (By similarity).
 CC -SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC
 CC EMBL: AL445063; CAC11304.1; -
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR003395; SMC_N.
 DR Pfam: PF02463; SMC_N; 1.
 KW DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
 FT NE_BIND 30 37 ATP (By similarity).
 FT DOMAIN 151 744 COILED COIL (POTENTIAL).
 SQ SEQUENCE 896 AA; 103435 MW; 2C584D700CCDCB2 CRC64;
 Query Match 5.6%; Score 106.5; DB 1; Length 896;
 Best Local Similarity 20.5%; Pred. No. 4.2;
 Matches 94; Conservative 64; Mismatches 158; Indels 143; Gaps 22;

QY 10 SEQIEKNNHNCCKTTDADAKKCENTPLGE-DEVCVMPRRONLCVHYLTKLNDKSE-RD 67
 Db 114 NDYIEKNI--TRSKDVELNSVSKQGEDDLSGDPARRK-----KLDELILEIK 163
 QY 68 LREAF-----TKSAAETFLROYNN-----SKNVEDKILHRM 102
 Db 164 LEETIYVLKDVYDSLQAGISNLDYLSENERDRDRLRYQDYVAELSKQIDEEAIESDL 223
 QY 103 IPPEFFRSMEYTFGDRDCLDPTDISKTIADHDVTTAK---KRTAVFQKIGSKT--- 154
 Db 224 LRKKEASAEVNVASKEHLMIDATLKNMMSLSDENRREBEIRKIDGKLOEISGTERYN 283
 QY 155 --TNGKVLERE---GWMKEEYGLST-WKGMCLAL-----SYNTEKK----- 190
 Db 284 EITSSVYVASRERIRRYWIDKQGIIDYRKMLKNIDQOVSYEDNMKKAELQADHDQYEI 343
 QY 191 ----MDE-----GVRTYLMKYTYKNNDI---KEYLEEFASPPPL----- 223
 Db 344 MQRMOEIKHEDDLDTYESKYSLINEIQKKKKEEYRKQKXDLGDELISRTLGAFAN 403
 QY 224 -----RWTEWGEDFVKNRKKEVLSLKKKDCSCLRNNGTSNKTCDNENCGA 271
 Db 404 ASELVAIYEIRIDIDEINTD-LGNLQVIGALRQKEE--IIRN--NMMEGHKKCPV 457
 QY 272 CKTQCEKTKKWMERMKKHYSS-----QKKPOLYNSATYNNGLAV 312
 Db 458 CGT--DLGDSGRRIREHYSSEDLINLETIDILHERASAIDEKILMESYLAKG--- 512
 QY 313 KEANSETYKN----PPEVTEANSAAKHADYLTQLENN 346
 Db 513 KIREYETIDROMKDLAQITDDENSLSITAYKRTKEQL 551

RESULT 13
 ID TOPL_CRIGR STANDARD; PRT; 767 AA.
 AC 007050;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA topoisomerase I (EC 5.99.1.2).
 GN TOPI OR TOP-1.
 OS Cricetus griseus (Chinese hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Cricetulus.
 CC NCBI_TaxID=10029;
 CC
 CC SEQUENCE FROM N.A.
 RX MEDLINE=94064611; PubMed=8244980;
 RA Tanigawa A., Bertrand R., Kohlhaagen G., Tabuchi A., Jenkins J.,
 RA Pommer Y.;
 RT "Cloning of Chinese hamster DNA topoisomerase I cDNA and
 RT identification of a single point mutation responsible for
 RT camptothecin resistance";
 RT J. Biol. Chem. 268:25463-25468(1993).
 CC
 CC -FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE
 CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.

CC -1- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
CC DNA, followed by passage and rejoining.
CC -1- ENZYME REGULATION: SPECIFICALLY INHIBITED BY CAMPOTHECIN (CPT), A
CC PLANT ALKALOID WITH ANTITUMOR ACTIVITY.
CC -1- SUBUNIT: MONOMER.
CC -1- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
CC RELAX ONLY NEGATIVE SUPERCOILS.
CC -1- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA
CC BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN
CC WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS
CC AT ONE END OF THE ENZYME-SEVERED DNA STRAND.
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC TYPE I TOPOISOMERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: Z21624; CAA/9747.1; -
DR EMBL: Z21625; CAA/9748.1; -
DR PIR: S32697; S32697.
DR HSSP: P11387; 1A35.
DR Interpro: IPR001631; Topoisomerase_I.
DR Pfam: PF01028; Topoisomerase_I.
DR Pfam: PF02919; Topoisomerase_I_N.
DR PRINTS: PR00416; EUTPISMRASE1.
DR SMART: SM00435; TOPEUC; 1.
DR PROSITE: PS00176; TOPOISOMERASE_I_EUK; 1.
KM Isomerase: Topoisomerase: DNA-binding.
FT ACT SITE 725 725 DNA CLEAVAGE (BY SIMILARITY).
FT VARIANT 505 505 S -> G (IN CPT-RESISTANT CELL).
SQ SEQUENCE 767 AA; 90867 MW; 34736D424EF35A9 CRC64;

Query Match
Best Local Similarity 20.5%; Score 106; DB 1; Length 767;
Matches 69; Conservative 47; Mismatches 106; Indels 114; Gaps 17;

OY 14 EKNINHNCKTQEDAKWCKENTKLGEDGVCMPPRRONLCVHYLTRLNDSDKEEDLREAFI 73
DB 89 EKRRKEXIRASGDGAKIRKE-----KENGFSPPR-----IKDPDDG----- 126
OY 74 KSAAEFTLLROYYSKKNVEDDKILHRDMIPPEFRSMYTTGDRDCLDT-DISE--- 129
DB 127 -----YAPPKEDIKPLKRPDEDD-----ADYKPKKIKTEDIKKEKK 164
OY 130 -KIADHDVTTAKKITTAVFQIGSKTTNGKVLERE-----GMWKE-YGLSI-W----- 175
DB 165 KRLDEEDGCKLKTNNKDKKAGSDNKKKPKKEEDGKWKWWEDEKRPEDIKKLEH 224
OY 176 KQMLCALSYNETFKMKDEGVRTY---LMKYLYKNNDIKEYL-----EEFASRPFLR-W 225
DB 225 KQPVAPRPY-----EPLPEGVKRYDGKVKLSPKAEVATFPKMLDHEYTTKEIFRKNF 280
OY 226 VTEMEDPYKNNKKKELVSLKKKCDSCITLRNNGTSNKTCDNENCGACKQCKEYKKNMER 285
DB 281 FYDMRKEMTNDKKNVITNL-SKCDP-----TOMSOYERKDOSE 316
OY 286 WKHHYSOKKKPOLYKNSATYNNGLAVKENSETYK 321
DB 317 ARKQMSKEK-----LKIKENKELK 338

RESULT 14
RBPL_PLAVB
AC 000738; STANDARD: PRT; 2869 AA.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Reticulocyte binding protein 1 precursor.
GN RBPL.
OS Plasmodium vivax (strain Belen).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
ON [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
RT merzoites."
RL Cell 69:1213-1226(1992).
CC -1- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
CC HUMAN RETICULOCYTE CELLS.
CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Membrane-bound.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: M88097; AAA29743.1; -
KM Malaria: Receptor; Signal; Transmembrane.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 2869 RETICULOCYTE BINDING PROTEIN 1.
FT DOMAIN 18 2807 EXTRACELLULAR.
FT TRANSMEM 2808 2826 POTENTIAL.
FT DOMAIN 2827 2869 CYTOPLASMIC.
FT SITE 1030 1032 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 2559 2601 CELL ATTACHMENT SITE (POTENTIAL).
SQ SEQUENCE 2869 AA; 330213 MW; B9DBE44205EBCF CRC64;

Query Match
Best Local Similarity 21.6%; Score 105.5; DB 1; Length 2869;
Matches 85; Conservative 47; Mismatches 123; Indels 139; Gaps 19;

OY 17 NINHNCK-TEDAKWKCKENTKLGEDGVCMPPRRONLCVHYLT-----KL 59
DB 998 SINCKKYNTDIDLRSKIKITLREYQKEMPRKDCGENTYALLSLDKMGKINEKL 1057
OY 60 ND-----DSKEDLEAFIKSAAEFTLLROYYSKKNVEDDKT-LHRMIPPEFRSMF 112
DB 1058 NGRNLNSIDTKED-----LKEYSES-----KHLSDQKGPQ--DPL 1096
OY 113 YTFGDRDCLDTD-----ISEKIADHDVTTAKKITTAVFQIGSK-----TTNGKVV 160
DB 1097 NTDIEMEDIKRVDENLVNVOYISENKKYTLFKNNSTVYLEAMHSIHNTYVAHGITSKNEI 1156
OY 161 LEREGWKKYGLSIWKGMLCALSYNETFKMKDEGVRTYLMKYLYKNNDIKEYLEEFASRP 220
DB 1157 LKSV-----KEV-----EDKLNLEQNEDEYKVV-----KNPENKOLEAIRGSM 1195
OY 221 PLRLWTEGDEGFVNNRKKELVSLKKKCDSCITLRNNGTSNKTCDNENCGACKTQC----- 276
DB 1196 SKLK-----EVINKHSEMTQLESTAN--TLKSNAKKEREHDEELNKTGQWRDIY 1246
OY 277 EYKKWMMERKKHYSOKKKPOLYKNSATYNNGLAVKENSETYKNDPE----- 325
DB 1247 EKLKTAELKE-----GTVNE-----LKDNENKANKYERPEPENNIIIGHVLE 1288
OY 326 -----VTEANSAKHARDYIKTOLENNI 347
DB 1289 RITVEKDKAGKYVEEMNS-----LKTQIKELKI 1315

RESULT 15
ENPL_CATRO
ID ENPL_CATRO STANDARD: PRT; 817 AA.
AC P35016;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2003, 14:54:09 : Search time 54.8604 Seconds
(without alignments)
1318.301 Million cell updates/sec

Title: US-10-087-013-9
Perfect score: 1913
Sequence: 1 PCKMYOKLISEQIEKNINHN.....AKHARDYKLTQENNICITNG 351

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1913	100.0	1685	5	Q9U4A2	Q9U4A2 plasmodium
2	646	33.8	921	5	Q25989	Q25989 plasmodium
3	629.5	32.9	2658	5	Q8T5G0	Q8T5G0 plasmodium
4	626.5	32.7	2527	5	Q95W83	Q95W83 plasmodium
5	625.5	32.7	3026	5	Q26031	Q26031 plasmodium
6	562.5	29.4	3026	5	Q26030	Q26030 plasmodium
7	550.5	28.8	3542	5	Q9U5M2	Q9U5M2 plasmodium
8	547.5	28.6	1711	5	Q96108	Q96108 plasmodium
9	520	27.2	2706	5	Q15870	Q15870 plasmodium
10	501	26.2	3006	5	Q26032	Q26032 plasmodium
11	500	26.1	3287	5	Q8T326	Q8T326 plasmodium
12	495.5	25.9	1615	5	Q8T325	Q8T325 plasmodium
13	460.5	24.1	2710	5	Q9XZB8	Q9XZB8 plasmodium
14	444.5	23.2	2212	5	Q94657	Q94657 plasmodium
15	434.5	22.7	2169	5	Q97312	Q97312 plasmodium
16	433.5	22.7	2042	5	Q25766	Q25766 plasmodium

17	432	22.6	5 Q25733	Q25733 plasmodium
18	429.5	22.5	5 Q9U0G6	Q9U0G6 plasmodium
19	418	21.9	5 Q9B3F0	Q9B3F0 plasmodium
20	403.5	21.1	5 Q97324	Q97324 plasmodium
21	403.5	21.1	5 P90580	P90580 plasmodium
22	402	21.0	5 Q9B3E6	Q9B3E6 plasmodium
23	399.5	20.9	5 Q26033	Q26033 plasmodium
24	398.5	20.8	5 Q96296	Q96296 plasmodium
25	397.5	20.8	5 Q61077	Q61077 plasmodium
26	384.5	20.1	5 Q9NFB6	Q9NFB6 plasmodium
27	375	19.6	5 Q60991	Q60991 plasmodium
28	374.5	19.6	5 Q26034	Q26034 plasmodium
29	364	19.0	5 Q00832	Q00832 plasmodium
30	360.5	18.8	5 Q25734	Q25734 plasmodium
31	355.5	18.6	5 Q9U0G5	Q9U0G5 plasmodium
32	353	18.5	5 Q25987	Q25987 plasmodium
33	350	18.3	5 Q9U4X0	Q9U4X0 plasmodium
34	345	18.0	5 Q9B3E7	Q9B3E7 plasmodium
35	344	18.0	5 Q9B3E9	Q9B3E9 plasmodium
36	336.5	17.6	5 Q00831	Q00831 plasmodium
37	333.5	17.4	5 Q8T9N3	Q8T9N3 plasmodium
38	333.5	17.4	5 Q8T9L8	Q8T9L8 plasmodium
39	333.5	17.4	5 Q96724	Q96724 plasmodium
40	333.5	17.4	5 Q8W531	Q8W531 plasmodium
41	329	17.2	5 Q27361	Q27361 plasmodium
42	329	17.2	5 Q25842	Q25842 plasmodium
43	325	17.0	5 Q25739	Q25739 plasmodium
44	325	17.0	5 P90585	P90585 plasmodium
45	323	16.9	5 Q27390	Q27390 plasmodium

ALIGNMENTS

RESULT 1	
ID Q9U4A2	PRELIMINARY; PRT; 1685 AA.
AC Q9U4A2;	
DT 01-MAY-2000 (TREMBLrel. 13, Created)	
DE 01-MAY-2000 (TREMBLrel. 13, Last sequence update)	
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)	
DE Variant surface protein PfEMP1 (Fragment).	
GN VAR.	
OS Plasmodium falciparum.	
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
OX NCBI_TaxID=5833;	
RN (1)	
RP SEQUENCE FROM N.A.	
RC STRAIN-IT;	
RX MEDLINE=20144115; PubMed=10677532;	
RA Smith J.D., Grais A.G., Krieck N., Hudson-Taylor D., Kyes S., Fagen T., Pinches R., Baruch D.I., Newbold C.I., Miller L.H.;	
RT "Identification of a Plasmodium falciparum intercellular adhesion molecule-1 binding domain: A parasite adhesion trait implicated in cerebral malaria."	
RT Proc. Natl. Acad. Sci. U.S.A. 97:1766-1771(2000).	
RL EMBL: AF193424; AAF18980.1;	
DR InterPro: IPR000454; Eub.ATPase_Csub.	
DR InterPro: IPR004258; PFEMP.	
DR Pfam: PF03011; PFEMP. 1.	
DR PROSITE: PS00605; ATPASE_C; 1.	
FT NON_TER 1685 1685	
SO SEQUENCE 1685 AA; 193671 MW; D1FDA2666B0551E CRC64;	
Query Match	100.0%; Score 1913; DB 5; Length 1685;
Best Local Similarity	100.0%; Pred. No. 7,1e-124;
Matches 351; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PCKMYOKLISEQIEKNINHNCKTTEADAKKCKCENTKIGEBGVCVMPRRONLCVHYLTQKN 60	
DB 1224 PCKMYOKLISEQIEKNINHNCKTTEADAKKCKCENTKIGEBGVCVMPRRONLCVHYLTQKN 1283	
QY 61 DSKKEEDLREAFIKSAAEFTLLRQYNSKNVNDKILHRDMIPPEFRSMFYTFGDYRD 120	

```

Db 1284 DSKNEEDLEAFIKSAAEFTLLROYNSKNVEDKILHRDMPPEFRSMFTGYD 1343
Qy 121 ICLDPTDISEKTIADHDVTTAKKKTAVFQKISGKTNGKAKYEREGWKEYGISTMKMLC 180
Db 1344 ICLDPTDISEKTIADHDVTTAKKKTAVFQKISGKTNGKAKYEREGWKEYGISTMKMLC 1403
Qy 181 ALSTYETKKNDEGVRYTLKMYIKYKNDIKKEYLEEFASRPPLRWVTEMGDEFPVNRKKE 240
Db 1404 ALSTYETKKNDEGVRYTLKMYIKYKNDIKKEYLEEFASRPPLRWVTEMGDEFPVNRKKE 1463
Qy 241 LVSLAKKCDSCGLRNNGTSNKTCDNENCGACKTQCEKYYKMMERKKHYSSQKKPOLY 300
Db 1464 LVSLAKKCDSCGLRNNGTSNKTCDNENCGACKTQCEKYYKMMERKKHYSSQKKPOLY 1523
Qy 301 KNSATYNNGLAVKEANSFTYKNDPEVTEANSKAHARDYLKTQLENMCTG 351
Db 1524 KNSATYNNGLAVKEANSFTYKNDPEVTEANSKAHARDYLKTQLENMCTG 1574

```

OUT 2

```

ID 025989 PRELIMINARY; PRT; 921 AA.
AC 025989;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE (clone pS1H) ORF (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRATN-JDP8;
RA MEDLINE=95350219; PubMed=7624377;
RA Peterson D.S., Miller L.H., Welles T.E.;
RT "Isolation of multiple sequences from the Plasmodium falciparum genome
RT that encode conserved domains homologous to those in erythrocyte-
RT binding proteins."
RL Proc. Natl. Acad. Sci. U.S.A. 92:7100-7104 (1995).
DR EMBL: L38454; AACC37240.1; -.
DR InterPro: IPR004258; PfEMP.
DR Pfam: PF03011; PfEMP; 1.
FT NON_TER 1
SQ SEQUENCE 921 AA; 104711 MW; D98BC07763628746 CRC64;

```

Query Match 33.8%; Score 646; DB 5; Length 921;
 Best Local Similarity 41.0%; Pred. No. 1.3e-36;
 Lches 150; Conservative 54; Mismatches 112; Indels 50; Gaps 16;

```

2 CKWVKLISEQIEKNNIHNCKTED-AKKCENT-KLGEDEGVCMPPRRONLCVHYTLK 58
Db 396 CEIYAEKLLKKNRTTYGECYRKETISEWTCDSKTKMGO-HGACIPPRQKCLHYLEK 454
Qy 59 LINDSKBEDLEAFIKSAAEFTLLROY--NSKVED-DKILHDMIPPEFRSMFTY 114
Db 455 IMNTNED-LKYAFICAAAEFTLLWQNYKKDKNGAEDLDKLGIIIPEDRKQMYFT 512
Qy 115 FGDRICIDPTDISEKTIADHDVTTAKKKTAVFQKISGKTNGKAKYEREGWKEYGIS 173
Db 513 FGDRICIDPTDISEKTIADHDVTTAKKKTAVFQKISGKTNGKAKYEREGWKEYGIS 173
Qy 174 IMKGMICALSYNT-----ET-KKMDGVRYTLKMYIKYKNDIKKEYLEEFASRPPLRW 225
Db 567 IMKGMICALSYNTSLANNVNPETHKKTLEEGNNNF-EKVI-F-GSSSTTLKFSRPOFLRW 624
Qy 226 VTEMGDEFPVNRKKEVLVSLAKKCDSCGLRNNGTSNKTCDNENCGACKTQCEKYYKMMER 285
Db 625 LVEWGNFCEQKKEKVKVLLAKCKDCDDVDGDC--KC--NGKCVACKQOCQYHSMIGI 678
Qy 678 WKHYSSQKKPOLYKNSATYNNGLAVKEANSFTYKNDPEVTEANSKAHARDYLKTQLEN 345
Db 679 WKHYSSQKKPOLYKNSATYNNGLAVKEANSFTYKNDPEVTEANSKAHARDYLKTQLEN 345

```

Qy 346 MICTG 351
 Db 722 MKCVNG 727

RESULT 3

```

ID 08T560 PRELIMINARY; PRT; 2658 AA.
AC 08T560;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Erythrocyte membrane protein-1 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRATN-JDP8;
RA Chattopadhyay R., Pillai C.R., Chitnis C.;
RT "Identification of a domain responsible for binding to intercalular
RT adhesion molecule-1 from a Plasmodium falciparum field isolate."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY028643; AAK49742.1; -.
FT NON_TER 2658
SQ SEQUENCE 2658 AA; 301326 MW; ABE33BC9167E7B7 CRC64;

```

Query Match 32.9%; Score 629.5; DB 5; Length 2658;
 Best Local Similarity 39.4%; Pred. No. 6e-35;
 Matches 146; Conservative 59; Mismatches 107; Indels 59; Gaps 17;

```

2 CKWVKLISEQIEKNNIHNCKTEDA-KKCENTKLGEDEGVCMPPRRONLCVHYTLK 54
Db 1307 CKTVNDILSTDRTKQYGCMAKIKININISYPMWTCVNSKFENNEGPMPPRRQKCLCY 1366
Qy 55 YLTKLNDKSEEDLEAFIKSAAEFTLLROYNSKNVEDKILHRDMPPEFRSMFTY 114
Db 1367 YLTKLNDKSEEDLEAFIKSAAEFTLLROYNSKNVEDKILHRDMPPEFRSMFTY 1426
Qy 115 FGDRICIDPTDISEKTIADHDVTTAKKKTAVFQKISGKTNGKAKYEREGWKEYGIS 174
Db 1427 YGDRICIDPTDISEKTIADHDVTTAKKKTAVFQKISGKTNGKAKYEREGWKEYGIS 1476
Qy 175 WKGMICALSYNT-----ET-KKMDGVRYTLKMYIKYKNDIKKEYLEEFASRPPLRW 228
Db 1477 WKGMICALSYNTSLANNVNPETHKKTLEEGNNNF-EKVI-F-GSSSTTLKFSRPOFLRW 624
Qy 229 WGEDEFPVNRKKEVLVSLAKKCDSCGLRNNGTSNKTCDNENCGACKTQCEKYYKMMER 287
Db 1531 WGEDEFPVNRKKEVLVSLAKKCDSCGLRNNGTSNKTCDNENCGACKTQCEKYYKMMER 1586
Qy 288 KHYSSQKKPOLYKNSATYNNGLAVKEANSFTYKNDPEVTEANSKAHARDYLKTQLEN 345
Db 1587 KHYSSQKKPOLYKNSATYNNGLAVKEANSFTYKNDPEVTEANSKAHARDYLKTQLEN 1637
Qy 342 ANHARDYLKTQ 342
Db 1638 YSKAGYIKKE 1648

```

RESULT 4

```

ID 095W83 PRELIMINARY; PRT; 2527 AA.
AC 095W83;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Erythrocyte membrane protein 1 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

```

OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21442075; PubMed=11557894;
 RA Flick R., Scholander C., Chen Q., Fernandez V., Pouvelle B., Gysin J.,
 Wahlgren M.,
 RT "Role of nonimmune IgG bound to PfEMP1 in placental malaria."
 RL Science 293:2098-2100(2001).
 DR EMBL: AF365677; AAL12845.1;
 DR InterPro: IPR004258; PfEMP.
 DR Pfam: PF03011; PfEMP. 2.
 FT NON_TER 2527 2527
 SO SEQUENCE 2527 AA; 292137 MW; 59EE2766BF0425E4 CRC64;
 Query Match 32.7%; Score 626.5; DB 5; Length 2527;
 Best Local Similarity 37.3%; Pred. No. 9.2e-35;
 Matches 139; Conservative 59; Mismatches 124; Indels 51; Gaps 11;
 1 PCKWVQKLISEQIEKNNIHC-KKTEDAK-----WK-C--ENTKLGEDGVCPRRQNIC 52
 1239 PCNVVKTLLKDKNENSDIGECIRKYKDNENKYPWDCRNKIKIG-EGAVMPPRQKLC 1297
 OY 53 VHYLTKLNDKSEEDIRAFIKSAAEFFLLROYN-----SKNVEDKTLH 99
 DB 1298 VDFLKLKNDQ-TDKLRDAFIKSAAEFFLLSMHRYKEDKKENPTEAPSLDVEAQTOLN 1356
 OY 100 RDMIPPEFRSMFTFGDRIKLDPTDISEKADHDVTTAKKITAVQXIGSKTTNGK 159
 DB 1357 DGIIPERKROMFTFGDRIKLDGIG-----NDPIVNNNTITAFQNGAQNPSDDT 1411
 OY 160 VLEBEGWKKEXGLSIWGMICALSYNETKMKDEGVTLYLKNNDIKYLEEFASR 219
 DB 1412 DSOQOVFGTGYKDIWEGMCALEKITEKQMTFQDKHETFDKYTFPSDKSPILKEFQOT 1471
 OY 220 PPLPLRWYENGEDEVYKRNKKELVLSKKKCDSTLRNNTSKTKTCDN-ENCGACKTQCEK 278
 DB 1472 PQLRMMTEWHEHCEREKREVEILRTFCACCEVSDNGS---CHKNEECOKREECK 1528
 OY 279 YKKMERKKHYSSOKKKFOLYKNSATYNNGLAVKANSEYKNDPEVTEANSKAKHARDY 338
 DB 1529 YOKMLEKKRENYDKOKYF-----KDKEGINDPDTIOST---EAYEY 1569
 OY 339 LKTQLENNMCTNG 351
 DB 1570 LGGKLLKNTTFTSG 1582
 OY 331
 ID 026031 PRELIMINARY; PRT; 3078 AA.
 AC 026031;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Variant-specific surface protein.
 GN VAR-1.
 OS Plasmodium falciparum (isolate Dd2).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=57267;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DD2;
 RX MEDLINE=95330813; PubMed=7606788;
 RA Su X.Z., Heathcote V.M., Wertheimer S.P., Guinet F., Herrfeldt J.A.,
 Peterson D.S., Ravetch J.A., Wellems T.E.;
 RT "The large diverse gene family var encodes proteins involved in
 cytoadherence and antigenic variation of Plasmodium falciparum-
 infected erythrocytes."
 RL Cell 82:89-100(1995).
 DR EMBL: L40608; AAA75396.1;
 DR InterPro: IPR004258; PfEMP.
 DR Pfam: PF03011; PfEMP. 2.
 SO SEQUENCE 3078 AA; 349297 MW; C8037C2BC3CCD7C3 CRC64;

Query Match 32.7%; Score 625.5; DB 5; Length 3078;
 Best Local Similarity 38.8%; Pred. No. 1.3e-34;
 Matches 137; Conservative 61; Mismatches 126; Indels 29; Gaps 9;
 2 CKWVQKLISEQIEKNNIHC-KKTEDAKKKITAVFQKIGSKTTNGKVVLEBEGWKKEXGLSIWK 57
 DB 1347 CKIVEKILEGKNGTGTWCECPNKESYPMDCRNNDISHDGACMPPRQKLCVLYIAHES 1406
 OY 58 KLNDKSEEDIRAFIKSAAEFFLLROYKSKNVEDKTLHROMIPPEFRSMFTFGD 117
 DB 1407 QTEMKTDDNLKDAFIKTAAEFFLLSMWYKSKNDESEKILLDRGILPSQILRSMWYTTGD 1466
 OY 118 YRDICLDPTDISEKADHDVTTAKKITAVFQKIGSKTTNGKVVLEBEGWKKEXGLSIWK 177
 DB 1467 YRDICLNDTISK--QNDVAAKKDKIGTFPSKDSKSSG---LSRQMMKNTNGPEIKWG 1521
 OY 178 MLCALS-YNETKMKDEGVTLYLKNNDIKYLEEFASRPFLLRWTEWGEDVYKN 236
 DB 1522 MLCALTRYVTDITDMKRRIKNDYSYDKVNOGNGNPSLEEFAPQFLRMWTEWGEDVYCAE 1581
 OY 237 RKELVLSKKKCDSTLRNNTSKTKTCDNENCGACKTQCEKYYKMMERKKHSSQKK 296
 DB 1582 ROKRENIITKAC-----NEINSTQCCNDARH--RCNQACRAYOEVENKKKEFSGQTNN 1633
 OY 297 FOLYKNSATYNNGLAVKANSEYKNDPEVTEANSKAKHARDYLTQLENNICT 349
 DB 1634 FVLKRVNPOD-----PEYKGYEKDGVOPIQGN-----EYLQKCDNNKCS 1675
 RESULT 6
 ID 026030 PRELIMINARY; PRT; 3026 AA.
 AC 026030;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Variant surface protein (Fragment).
 GN VAR.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE OF 2044-2922 FROM N.A.
 RC STRAIN-IT 4/25/5;
 RX MEDLINE=95330813; PubMed=7606788;
 RA Su X.Z., Heathcote V.M., Wertheimer S.P., Guinet F., Herrfeldt J.A.,
 Peterson D.S., Ravetch J.A., Wellems T.E.;
 RT "The large diverse gene family var encodes proteins involved in
 cytoadherence and antigenic variation of Plasmodium falciparum-
 infected erythrocytes."
 RL Cell 82:89-100(1995).
 RN [2]
 RP SEQUENCE OF 2044-2922 FROM N.A.
 RC STRAIN-IT 4/25/5;
 RA Smith J.D., Chitnis C.E., Craig A.G., Roberts D.J.,
 Hudson-Taylor D.E., Peterson D.S., Plinches R., Newbold C.I.,
 Miller L.H.;
 RT "Switches in the expression of Plasmodium falciparum var genes
 correlate with changes in antigenic and cytoadherent phenotypes of
 infected erythrocytes."
 RL Cell 0:0-0(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-IT 4/25/5;
 RX MEDLINE=99094502;
 RA Smith J.D., Kyes S., Craig A.G., Fagan T., Hudson-Taylor D.,
 Miller L.H., Baruch A.I., Newbold C.I.;
 RT "Analysis of adhesive domains from the AMAV Plasmodium falciparum
 erythrocyte membrane protein-1 identifies a CD36 binding domain."
 RL Mol. Biochem. Parasitol. 97:133-148(1998).
 RN [4]
 RP SEQUENCE FROM N.A.

[illegible]

FT	NON-TER	3542	3542
SEQUENCE	3542	AA;	413089 MM; 970D85EE8BD42BC2 CRC64;
Query Match	28.8%;	Score 550.5;	DB 5; Length 3542;
Best Local Similarity	39.0%;	Pred. No. 2.4e-29;	
Matches 133;	Conservative 48;	Mismatches 115;	Indels 45; Gaps 14;
QY	2	CKMYQKLSBQLEIKNNIHC--KTEDA--KWKCENTKLGEDGVCMPFRNOLCVHILT	57
DB	1261	CTVYNDILKEDKNGKQYVEDCHPRKSNCSYPMQCGNINLVBDPRCMPRRQKCLVHFLA	1320
QY	58	KLNDSDK--BEDKREAFRTKSAAEFTLRLQYNSKNVED--DKILFROMIPPEFPFPM	111
DB	1331	NDNEKTKLQSOYNLKEAFITKSAAEFTFSWYIYKSKDEGNELDLEKEGILPPELFISM	1360
QY	112	FYTFGDPYDILHDIDISEKINDHVTYAKRKITFAFORIGSKTTNGKAVLREECGMMRYG	171
DB	1381	FYTFGDPYDILFEGDIDISK--GHGEGSKLKEIDISLFRKGDDKSPRGK---TRQEWIETHS	1435
QY	172	LSIMKGMICALSYNTETRTKMDGAVTYLIMKYIYN--NDIKLEYLEEFASRPPELRWTE	228
DB	1436	HEIMEDAMLCAL-VKIGAKKDD-----PLENGYNNVKKRSDDKSTLLEEFAPKRPQFLRWLTE	1469
QY	229	WGEDFVNRKKRKELYSLKAKKCDSTLRLNNGTSNKTODDNENCGA--CKTDCEKYKKMMERWK	287
DB	1490	WYDDYCYLRQYKYLKDVOEKCKS-----NDOLKCDFTGCNNKCEPQYVRYMKK-K	1535
QY	288	KHYSOSKKKPOLYKNSATYNN--GLAVKE--ANSERYKN	322
DB	1356	KWIPQDQYTYDDEDRKKRFROHIGVAVTDTGTINADYLN	1576

Query Match	Best Local Similarity	Matches	Conservative	Score	S47.5	DB 5	Length	1711
29	WKC---EWTKIGDEGVCMPPRRONLCVHYLTIKLNDSEKEDLRARAFKSAATFLLRQ	85	1:1	1:1	1:1	1:1	1:1	1:1
856	WRCIAPSGTISGKGALICVPRTEGLCYLYKELS-DTQKGLRAFAFKTAAGETYLWQ	914	1:1	1:1	1:1	1:1	1:1	1:1
86	YNSK-----NVEDDKT-LHRDMLPEPFPSMYFTGVDYRDLDTDISKIADH	134	1:1	1:1	1:1	1:1	1:1	1:1
915	KYRKDKONETASTELDIDDPQTQLNGGKITPEDFKRQMTYTFGDRDLFLGRIGY-----N	969	1:1	1:1	1:1	1:1	1:1	1:1
135	DVTTAKKTTAVPFGISGRTNGKV-LEREGWMEKYGSLWKGMLCALSYNTETKWDK	193	1:1	1:1	1:1	1:1	1:1	1:1

Db 970 DLKRVNNNTTAVPQN-GDHIPNGQKTRDROEFMGTYGKDIMVGMAL-----OEAG 1021
 QY 194 GVFETYLAKYTKKNDIKEY-----LEEFASRPPELVEMGEDFVKNRKKELVSKKK 248
 Db 1022 GKKTLETETYSVNTNGHITGLTKLNEFSSRPFLKMTNEMGOFCRERTTOLQILKERC 1081
 QY 249 DSCTLLNNGTSNKTCDNENCGACKTQCEKKKMMERKKHYSOSSOKKFFOLYKNSATYNN 308
 Db 1082 --WVYQYNGDKGK--DDKKE--KCTEACTYKEMLTNMDONKQKQNRITYEVAGTS---- 1131
 QY 309 GLAVKCANSETYKNDEPVEANSAKHARDYLKTQLENNICTNG 351
 Db 1132 -----PYKEDSDYKE---SKYAHGYLRKILKNIITCTSG 1161

RESULT 9

015870 PRELIMINARY; PRT; 2706 AA.
 015870;
 01-JAN-1998 (TREMBlrel. 05, Created)
 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PfEMP1 (Fragment).
 GN R29R-VAR1.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IT 4/25/5; PubMed=9230440;
 RX MEDLINE=97373957; PubMed=9230440;
 RA Rowe J.A., Moulds J.M., Newbold C.I., Miller L.H.;
 RT "Plasmodium falciparum rosetting is mediated by PfEMP1 and requires
 complement receptor 1";
 RL Nature 388:292-295(1997).
 DR EMBL: Y13402; CAI73831.1;
 DR EMBL: Y13403; CAI73831.1; JOINED.
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR004258; PFEMP.
 DR Pfam: PF03011; PFEMP, 2.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR NON_TER 2706 2706
 SQ SEQUENCE 2706 AA; 308162 MW; 1C33D55AD5317D68 CRC64;

Query Match 27.2%; Score 520; DB 5; Length 2706;

Best Local Similarity 35.0%; Pred. No. 2.3e-27;
 Matches 131; Conservative 61; Mismatches 118; Indels 64; Gaps 17;

2 CKNVOKLISEOIKNNIHNK--KTEDAKKCE--NTKLGEDEGVCMPPRRONLCVHYLT 58
 1
 Db 770 CALV-KGVLQOKNGSIDNCKNAKRNKKNEMQCKNTFVDNGEVCMPRRKKSICINHLT 828
 QY 59 LNDSEEDLEAFIKSAAEFTLL-RQYNSKNVEDDKILHDMIPPEFRSMFYFGD 117
 Db 829 EEOCTKKKYQLEAFIKSAKETNLMDKYKNDKN-EAEELLKKGKIPEDFMRTMIFYFGD 887
 QY 118 YRDICLDTISEKIDHDTTAKKITAVQXIGSKTTNGKVLREGGMRKEYGLSTWKG 177
 Db 888 FRPFCLENDGK-----DVQYKKNINKEV---NNSKRGFKKIDPENMNMENGPQIWWG 939
 QY 178 MLCALSYNETFKKMDGCVRTYLLKYYKNDIKEY-----LEEFASRP 221
 Db 940 MLCAL--IHDTKDS-----IKNDKNKYEVITLAKRDGSGNGMTLSFPAKKP 966
 QY 222 FLAVWTEMGEDFVKNRKKELVSLKKKDCSTLRNNGTSNKTCDNENCGACKTQCEYKK 281
 Db 987 FLVFWFVWYDCKEKEROKYLTVEASTCKSI---DGQLK-CD-----RCGNKNCDEYKK 1036
 QY 282 WMEWMKKHYSOKKRRPOLYKNSATYNN--GLAVKE--ANSETYKNDPE---VTEANS 331
 Db 1037 YMRKKKEWMLQDYKDKKRENGKIDGPIGITIYKDYLVANAEYLLKKKFTASCVTSSGK 1096

QY 332 AKH-ARDYLKTOLE 344
 Db 1097 AONSATEEVKKNIE 1110

RESULT 10

026032 PRELIMINARY; PRT; 3006 AA.
 ID Q26032
 AC Q26032;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Variant-specific surface protein.
 GN VAR-3.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FCR3;
 RX MEDLINE=95330813; PubMed=7606788;
 RA Su X.Z., Heatwole V.M., Wertheimer S.P., Guinet F., Herrfeldt J.A.,
 RA Peterson D.S., Ravetch J.A., Wellem T.E.;
 RT "The large diverse gene family var encodes proteins involved in
 cytoadherence and antigenic variation of Plasmodium falciparum-
 infected erythrocytes.";
 RL Cell 82:89-100(1995).
 DR EMBL: L40609; AAA75397.1;
 DR InterPro: IPR004258; PFEMP.
 DR Pfam: PF03011; PFEMP, 2.
 SQ SEQUENCE 3006 AA; 343769 MW; 8FD5FE475F943C74 CRC64;

Query Match 26.2%; Score 501; DB 5; Length 3006;

Best Local Similarity 30.0%; Pred. No. 5.3e-26;
 Matches 125; Conservative 63; Mismatches 123; Indels 106; Gaps 14;

1 PCKNVOKL-----ISEOIKNNIHNCKKTEDAKKCE--NTKLGEDEGVCMPPRRONLCV 54
 1
 Db 1755 PCOIVKTLFTTELELKACPTKYNGREKPPN-WKCISSGSASGSICTIPRRRKLYLH 1812
 QY 55 YL-----TKLNDSEK-----EEDLEAFIKSAAEFTLL-RQYNSKNVE----- 93
 Db 1813 KIEGVDTVSSDGETTPTIHDALREAFIOTAVETFEFLMHRKYKIKERKEBELONGTE 1872
 QY 94 -----DDKILHDMIPPEFRSMFYFGDYRDC-----IDFTDI 127
 Db 1873 LLPPAAQVSPEDNPEHPOKRLKRGKIPPEFRKQMFYLGDRDLCVGVKQDVAALEASG 1932
 QY 128 SERIADHDTTAKKITAVFQKIGSKTTNGK-----VLERGGWKKKEYGLSIWGMICALS 183
 Db 1933 DNKSGDKNINDISEKISVLEKSGEYTPPKPQGTTRKEEMWQKNGEIHWMAMICALT 1992
 QY 184 YNTEYTKKMDGCVRTYLL-----KYTKKNDIKEY-----LEEFASRP 222
 Db 1993 HNTDTRQVDVQVQGLFENGKNTPKNSQYQKNTTIVSSVNGCPIGNIKLEOFASRPTE 2052
 QY 223 LRWVTEMGEDFVKNRKKELVSLKKKDCSTLRNNGTSNKTCDNENCGACKTQCEYKK 281
 Db 2053 LRWLEEGEEFCRQYIKLERIKREC-----HKDNRKNCDDGFCCKEKCCKDQSSFE 2105
 QY 268 --NCGACKTQCEYKKMMERKKHYSOSSOKKRRPOLYKNSATYNNGLAVKANSETYKN 332
 Db 2106 TLKPCSAKSCSKYKWKISRKKBETLQKGAYEKQKADAGNNN-DYKEF-SKTLRN 2160

RESULT 11

087326 PRELIMINARY; PRT; 3287 AA.
 ID 087326
 AC 087326;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE PfEMP1 (Fragment).


```

Db      1019 IFPGTJSSCRVTKDTSQITKSLGDAQATTEKGDTHIDDKKLOE---WMTIHGPKIMEG 1075
Qy      178 MCALSYN--TEFKMDEGVRTYLMKYIK--NNDIE--YLEEFASRPPLRAVTEMGED 232
      1076 MCALTNNGSESEK-----KNLQDYSYNKLNAEKDDCCLEFASKPOFLRWYEMSD 1130
Qy      233 FVKNRKKEVLSTLKKDCSCLRNNGTSNKTCDNENCGACKTCQCEYKKMMERKKHYSS 292
      1131 FCEBERK---LEDKVEDVIC--AKDYECCKNNKNSGVCKKCKEYNTTKTKTOYES 1184
Qy      293 QKKKPOLYKNSATYNNGLAVKEANSETYKNDPEVTANSAKHARDYLKTOLENMICTG 351
      1185 QEGKF-----NTERKORCKPEY--NSYSKKDASEYLKDK-----CIDG 1219

RESULT 14
094657 PRELIMINARY; PRT; 2212 AA.
094657;
01-FEB-1997 (TEMBLrel. 02, Created)
01-FEB-1997 (TEMBLrel. 02, Last sequence update)
01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Erythrocyte membrane protein 1.
GN PEMPL.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JG;
RA Yang S.Q., Wollish W.S., Gut J., Wu J., Ahn J., Petersen C.,
RA Fujicka H., Aikawa M., Leach J.H., Nelson R.G.;
RT "The molecular cloning and DNA sequence analysis of Plasmodium
RT falciparum erythrocyte membrane protein 1."
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U31083; AAB06961.1;
DR InterPro; IPR004258; PFEMP.
DR Pfam; PFO3011; PFEMP; 2.
SQ SEQUENCE 2212 AA; 253594 MW; A33456648C652FC CRC64;

Query Match 23.2%; Score 443; DB 5; Length 2212;
Best Local Similarity 26.2%; Pred. No. 3.8e-22;
Matches 128; Conservative 59; Mismatches 133; Indels 168; Gaps 19;

Qy      8 LISQLEKNNIHCKTE-----DAKWC-----ENTKGED----- 39
      1272 IVAGVLTKELENACPTKYGPKAPTSMKCIPTKTYNAATGSESSGNGALQRAKRAVES 1331
Qy      40 -----EGCMPRRONLCVHYLTKLN-----DSKEE 66
      1332 GSPVTSNGSGICIPRRRL--YIOKLHWAGNIVVSGAQTPOGCTSSPSGKETPSD 1388
Qy      67 DUREAFIKSAAEFTLL--ROYNSKVE-----DD 95
      1389 KLTAFIAIOSAIEFTFLMDRYKKEKEKEKRVANGGLVPSLNGPPOQPGVGTGDSPOS 1448
Qy      96 KILHRMIPPEFRSMFYTGDRDCL--DIDISEKIDHVTAKKITAVPOKI-- 150
      1449 KLOQGTAVIPPELRQMFYTLGDVADIFEGKNDIVIDITKNGDKDIAEREKKIKADIEVLK 1508
Qy      151 --GSKTTNGKVLERGMKREYGLSIWGMICALSYNTERK-----KMDGVRTYL-- 199
      1509 NADSQPPSDK---RQTMWQNGEHITANGICALITYKEKDEKTPLKONGLSALMDEK 1565
Qy      200 -----MKYIY-----KND--IKYELEFASRPPLRWTEMGEDFYKNR 238
      1566 NKKPKDOKYOYDKVLKEDNSGTPKTDNHPPLPLNFISRPYFLIEWGTFCEK 1625
Qy      239 KELVSLKKKDCSTLNNNGTSNKTCDNEN-----CGACKTCCEKTKYK 282
      1626 KRLKIKVEK---MDEGKKOKCGSDGEDCEIRKODYSTVDFVCPCEGKYKREFYKRW 1681

```

```

Qy      283 MERKKHYSSOKKKPOLYKNSATYNNGLAVKEANSETYKNDPEVTANSAKHARDYLKTO 342
      1682 IEKKKDEYDKOKKAYNNQKTDARNN-----NDNAFS-----TTLDTCTTAGDFLQT- 1728
Qy      343 LENMICTN 350
      1729 LKNGPCKN 1736

RESULT 15
097312 PRELIMINARY; PRT; 2169 AA.
097312;
ID 097312;
AC 097312;
DT 01-MAY-1999 (TEMBLrel. 10, Created)
DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Var (3D7-varT3-2).
GN PFC1120C; MAL3P7.55.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=99376085; Pubmed=10448855;
RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
RA Churcher C.M., Craig A., Davies R.M., Devlin K., Felwell T.,
RA Gentles S., Gwilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
RA Horrocks P., Jagels K., Jassal B., Kyes S., McLean S., Moulé S.,
RA Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
RA Rutter S., Skelton J., Squares S., Sulston J.E.,
RA Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
RT falciparum."
RL Nature 400:532-538(1999).
DR EMBL; AL034559; CAB39061.1;
DR InterPro; IPR004258; PFEMP.
DR Pfam; PFO3011; PFEMP; 2.
SQ SEQUENCE 2169 AA; 246913 MW; 1D48ACB7AFAE3BE CRC64;

Query Match 22.7%; Score 434.5; DB 5; Length 2169;
Best Local Similarity 26.7%; Pred. No. 1.4e-21;
Matches 128; Conservative 50; Mismatches 127; Indels 175; Gaps 16;

Qy      1 PCKMYQKLSQLEKNNIHNC-----KTEDAKWC-----ENTLGEDEG----- 41
      827 PCDIVQTLF--ESTKNLSDACGLKYGGEKKEPPNMKCISSGSDTGSTTKQNDSESGEG 884
Qy      42 -----VCMPPRRONLCVHYLTKLNDSKEED----- 67
      885 HREKRTHESSDSTTSSGSVCVPRRRKLYVGLTIKAEARKSSTSPQEGSGVANASA 944
Qy      68 -----LREAFIKSAAEFTLLROYNSKVEDKI----- 97
      945 SSTSPPTDATQLRLDAFIAIOSAIEFTFLMKRYKAENTRONKSPGLNGVAQFGSYSGSE 1004
Qy      98 -----LHRMIPPEFRSMFYTGDRDCL--DIDISEKIDHVTAKKITAVPOKI-- 140
      1005 SEKTPQOQWQSGTTPDRLQMFYTLGDVADIFEGKNDIVIDITKNGDKDIAEREKKIKADIEVLK 1064
Qy      141 --KKTAVFOKIG--SKTTNGKVLERE--GMMKREYGLSIWGMICALSYNTERK-----K 189
      1065 ISEKIKVIEKSDTPSRTPGQPSNDPKSMWKHLAPITWGMIALYALTYDINTASGEK 1124
Qy      190 KMDGVRTY-----LMKYTYKNDIKE----- 211
      1125 KIEKDAVYKKLMDKANPKKKNQGOODYTYEYKVEIKEDSGQKASTASQTPSPRASGEN 1184
Qy      212 ---YLEEFASRPPLRWTEMGEDFYKNRKKEVLSTLKKDCSCLRNNGTSNKTCDNEN----- 265
      1185 KPTTLQSEVARKPYFYLEWGETFCEKERRKRLKIKVDCVEENFGARGTTKQKYSGD 1244
Qy      266 NENCG-----ACKTCCEKTKYKMMERKKHYSSOKKKPOLYKNSATYNNG 309

```

Db 1245 GEECSKIYENKDKIFKDLKPSKATPCGLYKRWIRKKDEYNKOKSAVNEQTK--YENG 1302

Search completed: June 20, 2003, 15:05:21
Job time : 56.8604 secs